Analysis of Defects in Transcriptional Regulation Mechanisms in High-Grade Serous Ovarian Cancer and Their Correlation with Clinic-Pathological Features

Thesis

Analysis of defects in transcriptional regulation mechanisms in high-grade serous ovarian cancer and their correlation with clinic-pathological features

Thesis submitted for the degree of Doctor of Philosophy

Open University, UK

Discipline of Life, Health and Chemical Sciences

Lara Paracchini

Master Degree in Biology Applied to Biomedical Research

Istituto di Ricerche Farmacologiche “Mario Negri” -IRCCS-, Milan, Italy

July 2020

Abstract

High grade serous epithelial ovarian cancer (HGS-EOC) is the most lethal gynaecological disease, usually diagnosed at advanced stages. Although most HGS-EOC patients are initially sensitive to therapy, about 20% of them does not respond to front-line platinum (Pt)-based chemotherapy, relapsing within 6 months from treatment. No molecular biomarkers are as yet available in the clinical practice to discriminate, at the time of diagnosis, the patients who will respond or not to first-line Pt treatment. The identification of a molecular signature associated with intrinsic Pt resistance would allow to select patients who will not respond to Pt compound, thus avoiding ineffective treatment and related toxicities.

In this thesis, defects in transcriptional regulation mechanisms associated to intrinsic Pt resistance were investigated using two different omic approaches – microarray and RNA sequencing. Moreover, considering the correlation between defects in DNA repair Homologous Recombination and Pt sensitivity, we have started to develop an academic assay which exploits high-throughput DNA sequencing to evaluate Homologous Recombination Deficiency (HRD). Results obtained from three independent HGS-EOC cohorts (n=1080) indicate that the combined expression profile of three genes – PRKG1, SDF2L1 and PPP1R12A – could represent an independent prognostic biomarker for Pt response and survival. In addition, 25 transcripts, 16 of which are unknown or partially known variants, have been identified that are differentially expressed between Pt-sensitive and Pt-resistant patients. These results, once validated, integrated and associated with the HRD status, could identify a molecular signature applicable in the clinical setting to enable stratification of patients at time of diagnosis, allowing to choose the best therapeutic option.

Acknowledgements

A Roberto, Anita e Michele,

le mie rocce.

Alla vita

"La vita non è uno scherzo.
Prendila sul serio
come fa lo scoiattolo, ad esempio,
senza aspettarti nulla
dal di fuori o nell'al di là.
Non avrai altro da fare che vivere.

La vita non é uno scherzo.
Prendila sul serio
ma sul serio a tal punto
che messo contro un muro, ad esempio, le mani legate,
o dentro un laboratorio
col camice bianco e grandi occhiali,
tu muoia affinché vivano gli uomini
gli uomini di cui non conoscerai la faccia,
e morrai sapendo
che nulla è più bello, più vero della vita.

Prendila sul serio
ma sul serio a tal punto
che a settant'anni, ad esempio, pianterai degli ulivi
non perché restino ai tuoi figli
ma perché non crederai alla morte
pur temendola,
e la vita peserà di più sulla bilancia".

Nazim Hikmet,1948

On living

"Living is no laughing matter:
you must live with great seriousness
like a squirrel, for example—
I mean without looking for something beyond and above living,
I mean living must be your whole occupation.

Living is no laughing matter:
you must take it seriously,
so much so and to such a degree
that, for example, your hands tied behind your back,
your back to the wall,
or else in a laboratory
in your white coat and safety glasses,
you can die for people—
even for people whose faces you've never seen,
even though you know living
is the most real, the most beautiful thing.

I mean, you must take living so seriously
that even at seventy, for example, you'll plant olive trees—
and not for your children, either,
but because although you fear death you don't believe it,
because living, I mean, weighs heavier."

Nazim Hikmet,1948

Preface and declaration

The work described herein was performed at the Istituto di Ricerche Farmacologiche "Mario Negri"- IRCCS in Milan, Italy from 2015 to 2020.

The work was performed under the supervision of Dr. Sergio Marchini (director of the study), Dr. Luisa Diomede (third party monitor) and Prof. Robert Brown (external supervisor).

This thesis has not been submitted in whole or in part for a degree or diploma or other qualifications to any other university.

The experimental work described here was performed by myself, but I declare that, although I have followed the various steps of analyses, the statistic and bioinformatic analyses presented in this work of thesis have been performed by the team of Prof. Chiara Romualdi (Univerisity of Padova, Italy) and by Dr. Luca Beltrame, who is part of my laboratory group.

Table of contents

List of Tables

6. Appendix

List of figures

1. Introduction

2. Aims

3. Material and Methods

4. Results

Publications by the candidate

Papers as first/co-first author:

- **L. Paracchini***, L. Beltrame*, A. Inglese et al., Genome-wide copy number alterations in circulating tumor DNA as a novel biomarker for high grade serous ovarian cancer patients. Jul 2020, under review.
- **L. Paracchini***, C. Pesenti*, M. Delle Marchette, L. Beltrame, T. Bianchi, T. Grassi et al., Detection of TP53 clonal variants in Papanicolaou test samples collected up to 6 years prior to high-grade serous epithelial ovarian cancer diagnosis. JAMA Network Open. 2020 Jul 1; 3(7): e207566-e207566
- G. Benvenuto*, P. Todeschini*, **L. Paracchini***, E. Calura, R. Fruscio, C. Romani et al., Expression profiles of PRKG1, SDF2L1 and PPP1R12A are predictive and prognostic factors for therapy response and survival in high-grade serous ovarian cancer. International Journal of Cancer. 2020 Jul 15; 147(2): 565-574
- **L. Paracchini**, L. Beltrame, L. Boeri, F. Fusco, P. Caffarra, S. Marchini et al., Exome sequencing in an Italian family with Alzheimer's disease points to a role for seizure-related gene 6 (*SEZ6*) rare variant R615H. Alzheimer's research & therapy. 2018 Dec 1; 12(1): 106
- P. Todeschini*, E. Salviato*, **L. Paracchini*,** M. ferracin, M.Petrillo, L. Zanotti et al., Circulating miRNA landscape identifies miR-1246 as promising diagnostic biomarker in high-grade serous ovarian carcinoma: A validation across two independent cohorts. Cancer Letters. 2017 March 1; 388: 320-327
- P.Martini*, **L. Paracchini*,** G. Caratti, M. Mello-Grand, R. Fruscio, L. Beltrame et al., lncRNAs as novel indicators of patients' prognosis in stage I epithelial ovarian cancer:

A retrospective and multicentric study. Clinical Cancer Research. 2017 May 1; 23(9): 2356-2366

- **L. Paracchini***, L. Mannarino*, I. Craparotta, C. Romualdi, R. Fruscio, T. Grassi et al., Regional and temporal heterogeneity of epithelial ovarian cancer tumor biopsies: implications for therapeutic strategies. Oncotarget. 2016 Sept 9; 5

Other papers:

- E. Calura, R. Fruscio, **L. Paracchini**, E. Bignotti, A. Ravaggi, P. Martini et al., miRNA Landscape in Stage I Epithelial Ovarian Cancer Defines the Histotype Specificities. Clinical Cancer research. 2013 Aug 1; 19(15): 4114-4123
- R. Dossi, R. Frapolli, S. Di Giandomenico, **L. Paracchini**, F. Bozzi, S. Brich et al., Antiangiogenic activity of trabectedin in myxoid liposarcoma: Involvement of host TIMP-1 and TIMP-2 and tumor thrombospondin-1. International Journal of Cancer. 2014 Jun 10; 136(3): 721-729
- L. Beltrame, M. Di Marino, R. Fruscio, E. Calura. D. Chapman, L. Clivio, et al., Profiling cancer gene mutations in longitudinal epithelial ovarian cancer biopsies by targeted next-generation sequencing: a retrospective study. Annals of Oncology. 2015 Jul 1; 26(7): 1363-1371
- M. Petrillo, G.F. Zannoni, L. Beltrame, E. Martinelli, A. Di Feo, **L. Paracchini** et al., Identification of high-grade serous ovarian cancer miRNA species associated with survival and drug response in patients receiving neoadjuvant chemotherapy: a retrospective longitudinal analysis using matched tumor biopsies. Annals of Oncology. 2016 Apr 1; 27(4): 625-634

- F. Pischiutta, L. Brunelli, P. Romele, A. Silini, E. Sammali, **L. Paracchini,** et al., Protection of Brain Injury by Amniotic Mesenchymal Stromal Cell-Secreted Metabolites. Critical Care Medicine. 2016 Nov 1; 44(11): e1118-e1131
- E. Calura, **L. Paracchini**, R. Fruscio, A. Di Feo, A. Ravaggi, J. Perrone et al., A prognostic regulatory pathway in stage I epithelial ovarian cancer: new hints for the poor prognosis assessment. Annals of Oncology. 2016 Aug 1; 27(8): 1511-1519
- V. Iori, A. M. Iyer, T. Ravizza, L. Beltrame, **L. Paracchini**, S. Marchini et al., Blockade of the IL-1R1/TLR4 pathway mediates disease-modification therapeutic effects in a model of acquired epilepsy. Neurobiology of disease. 2017 March 1; 99: 12-23
- C. Belgiovine, E. Bello. M. Liguori, I. Craparotta, L. Mannarino, **L. Paracchini** et al., Lurbinectedin reduces tumour-associated macrophages and the inflammatory tumour microenvironment in preclinical models. British Journal of cancer. 2017 Aug; 117(5): 628-238
- L. Mannarino, **L. Paracchini**, I. Craparotta, M. Romano, S. Marchini, R. Gatta et al., A systems biology approach to investigate the mechanism of action of trabectedin in a model of myelomonocytic leukemia. Pharmacogenomics Journal.2018 Jan; 18(1): 58-63
- E. Calura, M. Ciciani, A. Sambugaro, **L. Paracchini**, G. Benvenuto, S. Milite et al., Transcriptional Characterization of Stage I Epithelial Ovarian Cancer: A Multicentric Study. Cells. 2019 Dec; 8(12): 1554

List of Abbreviation

OC: Ovarian Cancer

EOC: Epithelial Ovarian Cancer

FIGO: International Federation of Gynaecology and Obstetrics

TVS: Transvaginal Ultrasound

CA-125: Cancer Antigen-125 or Carbohydrate Antigen-125

MMR: MisMatch Repair

HGS: High-Grade Serous

LGS: Low Grade Serous

HPF: High-Power microscopic Fields

IHC: ImmunoHistoChemistry

HR: Homologous Recombination

Mb: Mega Bases

CT: Computed Tomography

MRI: Magnetic Resonance Imaging

Pt: PlaTinum

PDS: Primary Debulking Surgery

RT: Residual Tumour

NACT: NeoAdjuvant ChemoTherapy

IDS: Interval Debulking Surgery

FDA: Food and Drug Administration

CBDCA: Carboplatin

PAC: PAClitaxel

VEGF: Vascular Enothelial Growth Factor

PFS: Progression-Free Survival

OS: Overall Survival

PARP: Poly (ADP-Ribose) Polymerase

PARPi: Poly (ADP-Ribose) Polymerase inhibitor

HR: Hazard Ratio

CI: Confidence Interval

HRD: Homologous Recombination Deficiency

PLD: Pegylated Liposomal Doxorubicin

EMA: European Medical Agency

ET-743: Trabectedin

CDDP:cisplatin

NER: Nucleotide Excision Repair

FA: Fanconi Anemia

HR: Homologous Recombination

HRD: Homologous Recombination Deficiency

dsDNA: Double Strand DNA

CTR1: Copper TranspoRter 1

ABC: ATP Binding Cassette

MRP: Multi-Resistance associate Protein

cMOAT: Canalicular Multispecific Organic Anionic Transporter

GSH: Glutathione

MT: Metallothioneins

DSB: Double Strand Break

ssDNA: Single Strand DNA

HJ: Hollyday Junction

LOH: Loss of Heterezygosis

TAI: Telomeri Allelic Imbalances

LST: Large-Scale Transition

EMT: Epithelial to Mesenchymal Transition

TGFβ: Transforming Growth Factor β

PFI: Platinum Free Interval

Bp: Base Pairs

gDNA: Genomic DNA

lncRNA: Long Non-Coding RNA

miRNA: Micro RNA

RLE: Relative Long Expression

DE: Differentially Expressed

dNTP: DeoxyriboNucleotide TriPhosphates

tDNA: Tumour DNA

FDR: False Discovery Rate

Cy3: Cyanine 3

logFC: log Fold Change

DEG: Differentially Expressed Genes

DET: Differentially Expressed Transcripts

AF: Allelic Fraction

lincRNA: Long Intragenic Non-Coding RNA

1.Introduction

1.1 Malignant Epithelial Ovarian Cancer

1.1.1 Epidemiology and risk factor

Ovarian cancer (OC) is the 6th more lethal cancer among females representing 4,4% of cancer-related deaths among women (1). The American Cancer Society has estimated approximately 21,750 new cases and 13,940 deaths in the USA in 2020. The proportion between new OC cases/year and deaths due to OC/year highlights the aggressiveness of this kind of tumour, and support the notion that OC can be considered as the most lethal gynaecological cancer among women (2).

Epithelial Ovarian Cancer (EOC) is the most common type of ovarian tumour, making up 90% of all primary ovarian cancer and it is considered as an age-related disease (median age of disease onset 65 years), as in sporadic forms its incidence increases in post-menopausal women (3). Despite ovarian cancer survival has almost doubled in the last 40 years in UK (4) due to the evolution of modern medicine, including improvements in surgery procedures, the overall 5-years survival rate still remains low and ranges between 30%-40%. This high mortality rate is due to many different factors that will be addressed only in part in this chapter. First, the majority of EOC (about 70%) is diagnosed when the disease is already metastatic, that means that it has already spread into the abdominal cavity (International Federation of Gynaecology and Obstetrics -FIGO- Stage III/IV): in these cases, the 5-years survival rate is about 29%, in comparison to 92% when the tumour is limited to the ovary (FIGO, stage I) (5). The tumour stage, which is a measure of the spread of the disease in the abdominal cavity is currently the main prognostic parameter (see paragraph 1.1.2).

Nowadays there are no effective strategies or screening tests for EOC early disease diagnosis. A prospective study made on a cohort of more than 100000 women in the United Kingdom (6) demonstrated transvaginal ultrasounds scans (TVS) and the measurements of CA-125 serum

protein levels, two assays routinely used in clinic for EOC screening, have a low positive predictive value, being unable to differentiate benign from malignant EOC in the case of TVS (7), or have low specificity and sensitivity in the case of CA-125 (8).

The strongest risk factor for EOC is constituted by hereditable genetic conditions, which are responsible of 15%-20% for EOC cases (9).

Among these hereditable syndromes the most common (65%-75%) is represented by the *Breast-Ovarian cancer syndrome*, which is an autosomal dominant genetic syndrome that increases the risk to develop EOC (high-grade serous histotype, see paragraph 1.2) and breast cancer during the life-time. This syndrome is due to germline mutations that lead to loss of function of two proteins codified by *BRCA1* (chromosome 17q21) and *BRCA2* (chromosome 13q12-13) genes. In these patients the cumulative EOC risk is 44% and 17% for patients with germline mutation in *BRCA1* and *BRCA2* genes respectively (10) instead of the 1.5% of the general population (11).

Another genetic syndrome that accounts for the 10-15% of hereditary EOC, is the *Lynch syndrome* that is typically associated with endometrioid and clear cell histotype (see paragraph 1.1.2). *Lynch syndrome* is a genetic autosomal dominant syndrome regarding germline mutations in Mismatch Repair (MMR) genes (*MLH1*, *MSH2* and *MSH6*), that is associated with an increased risk of colon-rectal, endometrial, breast, urinary tract, small bowel and EOC. In these patients the cumulative risk to develope EOC within 70 years is about 20% for MLH1 and MSH2 germline mutation carriers (11).

Other genetic syndromes associated with hereditary EOC include Peutz-Jegher (germline mutations in *STK11/LKB1* genes) and other rare disorders, such as Gorlin syndrome (germline mutations in *PTCH* gene) but they will not be discussed because out of the scope of the thesis.

Other non-genetic EOC risk factors (9) include:

- the number of lifetime ovulations (early age of menarche, late age of menopause, absence of pregnancy). The hypothesis of the "incessant ovulation" in fact, points that each ovulatory cycle is associated with an increase rate cellular division to repair the surface epithelium, leading to the arise of spontaneous DNA mutations
- smoking
- benign gynaecological conditions (i.e. polycystic ovary syndrome, endometriosis)
- use of talcum powder (12), although results about this point are controversial.

1.1.2 Classification

As previously state (paraghraph1.1.1) the term "EOC" is nowadays considered as a misleading term, as it does not refer to a single disease, but an heterogeneous subset of malignancies involving the ovary, characterized by different clinical, anatomo-pathological, etio-pathogenic, and molecular features (figure 1.1) (13) (14).

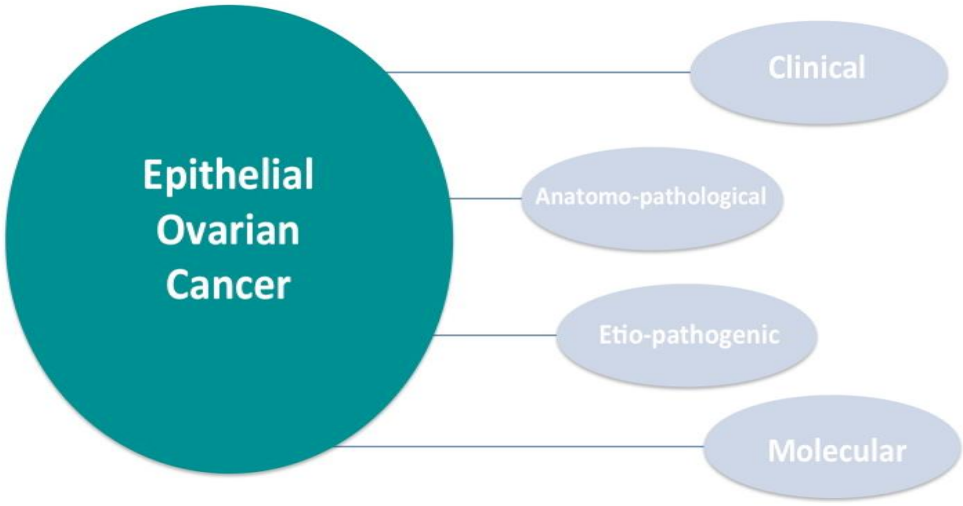


Figure 1.1 The heterogeneous nature of EOC considering clinical and biological aspects.

This extremely complex nature of EOC has led to different EOC classification systems that depend on the basis of the features considered (clinical, biological or genetics).

Independently from the histopathological classification, that identifies five main subtypes of EOC (high-grade serous –HGS-, low grade serous –LGS-, endometrioid, clear cells and mucinous), from surgical point of view EOC can be classified upon the extent and diffusion of the disease at time of diagnosis. The International Federation of Gynaecology and Obstetrics (FIGO) criteria identified four different stages and eight substages, the features of which are detailed in the table 1.1 (15).

Stage I: confinement to the ovary	
IA	Tumour limited to one ovary with an intact capsule or to Fallopian tubes.
IB	Tumour limited to both the ovaries with intact capsules or to both the Fallopian tubes.
IC	Tumour limited to one or both ovaries or Fallopian tubes, but with the capsule rupture or the presence of malignant cells on the epithelial surface. Malignant cells are also present in the ascites or peritoneal washing.
Stage II: involvement one or both ovaries or Fallopian tubes, with pelvic extension or primary peritoneal cancer	
IIA	Tumour is extended and/or implanted on the uterus and/or Fallopian tubes and/or ovaries.
IIB	Tumour has extended to other pelvic intraperitoneal tissues.
Stage III: involvement one or both ovaries or Fallopian tubes, or primary peritoneal cancer with spreading to the peritoneum outside the pelvis and/or metastasis to the retroperitoneal lymph nodes	
III A	Cytologically or histologically proven positive retroperitoneal lymph nodes only or microscopic extrapelvic peritoneal involvement
III B	Macroscopic peritoneal metastasis beyond the pelvis up to 2 cm diameter with or without metastasis to retroperitoneal lymph nodes
III C	Macroscopic peritoneal metastasis beyond the pelvis more than 2 cm diameter with or without metastasis to retroperitoneal lymph nodes
Stage IV: presence of distant metastasis, with malignant cells present in the pleural effusion and involvement of abdominal organs with parenchymal and extra-abdominal metastasis	

Table 1.1 International Federation of Gynaecology and Obstetrics (FIGO) staging criteria

Another type of EOC classification is based on the grade of tumour cell differentiation. Comparing the tissue and cancer cells to healthy ones, parameters as the tissue architecture, the degree of nuclei atypia and the mitotic index are considered.

On the basis of these parameters EOC can be classified in three different grades of differentiation ("three-tier grading"): G1, well differentiated; G2 moderately differentiated; G3 poorly differentiated. This "three-tier grading" could be not used for serous EOC, for which a "two-tier" grading system was adopted (high grade for poorly differentiated and low grade for well differentiated forms) (16).

Over the last years, with the development of more accurate and sensitive technologies to study genetic and genomic alterations that characterized EOC, new classification criteria have emerged. In particular the "dualistic model" proposed by Kurman and colleagues has been proposed to explain the pathogenesis and classified EOC on the basis of origin, clinical, biological and molecular features. According to this classification EOCs are subdivided into two groups named "type I" and "type II" tumours, the main characteristics of which are reported in table 1.2.(17).

Features	Type I	Type II
Stage	Frequently early stage	Almost always advanced stage
Tumour grade	Low grade	High grade
Histological subtypes	Serous, Clear Cells, Endometrioid, Mucinous, malignant Brenner tumour, seromucinous carcinoma	High Grade Serous EOC, undifferentiated carcinoma, carcinosarcoma
Proliferative activity	Generally low	Always high
Ascites	Rare	Common
Response to chemotherapy	Fair	Good (with recurrence within 18 months in the majority of cases)
Early detection	Possible	Challenging
Progression	Slow and indolent	Rapid and aggressive
Overall clinical outcome	Good	Poor
Chromosomal instability	Low	High
Mutations	*KRAS, BRAF, PTEN, CTNNB1, ERBB2, PI3CA, ARID1A*	*TP53*
Homologous recombination repair	Rarely defective	Frequently defective
Actionable mutations	Can be present	Rare

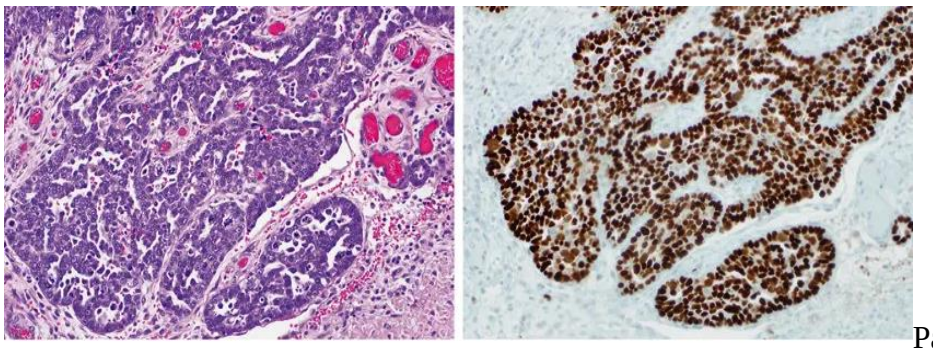
Table 1.2 Type I and type II EOC features

As reported in table above type II EOC is mainly characterised by most common and aggressive EOC histotype: the High Grade Serous (HGS) EOC. As HGS-EOC represents the focus of this thesis work, the other EOC histotypes are not considered in this context.

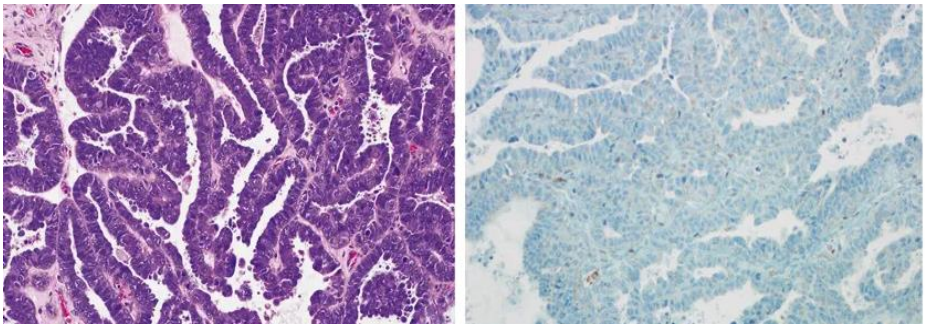
1.2 High-Grade Serous Epithelial Ovarian Cancer: clinical and molecular features

High-Grade Serous Epithelial Ovarian Cancer (HGS-EOC) is the most common type of EOC (about 80% of all EOC cases) and in 90% of cases it is diagnosed when tumour has already spread into the abdominal cavity and in intra- or extra- pelvic organs. The late diagnosis, the rapid growth and the aggressiveness of HGS-EOC explain the poor prognosis related to this specific EOC subtype (5-years survival rate about 30%).

From a histopathological point of view HGS-EOC is characterized by papillary or solid growth with slit-like glandular lumens pattern. Tumour cells appear with an intermediate size, with, prominent nucleoli and high mitotic index (activity greater than 12/10 high-power microscopic fields, HPF). A great numbers of immuno-histological biomarkers are used to better distinguish the HGS-EOC from the other EOC subtypes including WT1, ER, PR and PAX8 (14) but in particular the p53 protein. In fact, HGS-EOC is associated to an abnormal immunohistochemistry (IHC) staining for p53, that can be both overexpressed (figure 1.2 panel A) or completely absent (figure 1.2 panel B), depending on the nature of TP53 clonal mutation associated to the disease (18).



Panel A



Panel B

Figure 1.2 IHC staining for HGS-EOC. Panel A: high grade serous ovarian cancer with strong and diffuse immunoexpression of p53 (100% positive cells). Panel B: high grade serous ovarian cancer with completely absent immunoexpression of p53 (0% positive cells).

The abnormal ICH staining for p53 is due to the fact that about the 96% of HGS-EOC cases, is characterized by a clonal pathogenic mutation in *TP53* gene (19). In the remaining mutation negative HGS-EOC cases (4%), it has been demonstrated that the dysfunction of p53 pathway that characterized HGS-EOC, is related with chromosomal gain in *MDM2* and *MDM4*, which are involved in the degradation of p53 protein (20).

Somatic mutations in genes other than *TP53* occurs in no more than 5%-8% HGS-EOC cases, and they mainly involved *BRCA1*, *BRCA2*, *CDK12* and Homologous Recombination (HR) pathway-related genes (21).

As reported by Ciriello and collegues (22) few years ago in fact, the HGS-EOC is not characterized by high prevalence of somatic mutations (with the exception of clonal *TP53* alterations) or methylation events, but by multiple recurrent chromosomal gains and losses, which include the clonal somatic amplification in 8q24 and 3q24 genomic loci (23), *CCNE1* locus amplification (15% of HGS-EOC cases) that is mutually exclusive with *BRCA1/BRCA2*

gene mutations, and losses in important oncosoppressor genes as *NF1*, *RB1*, *PTEN* (figure 1.3 and figure 1.4) (9).

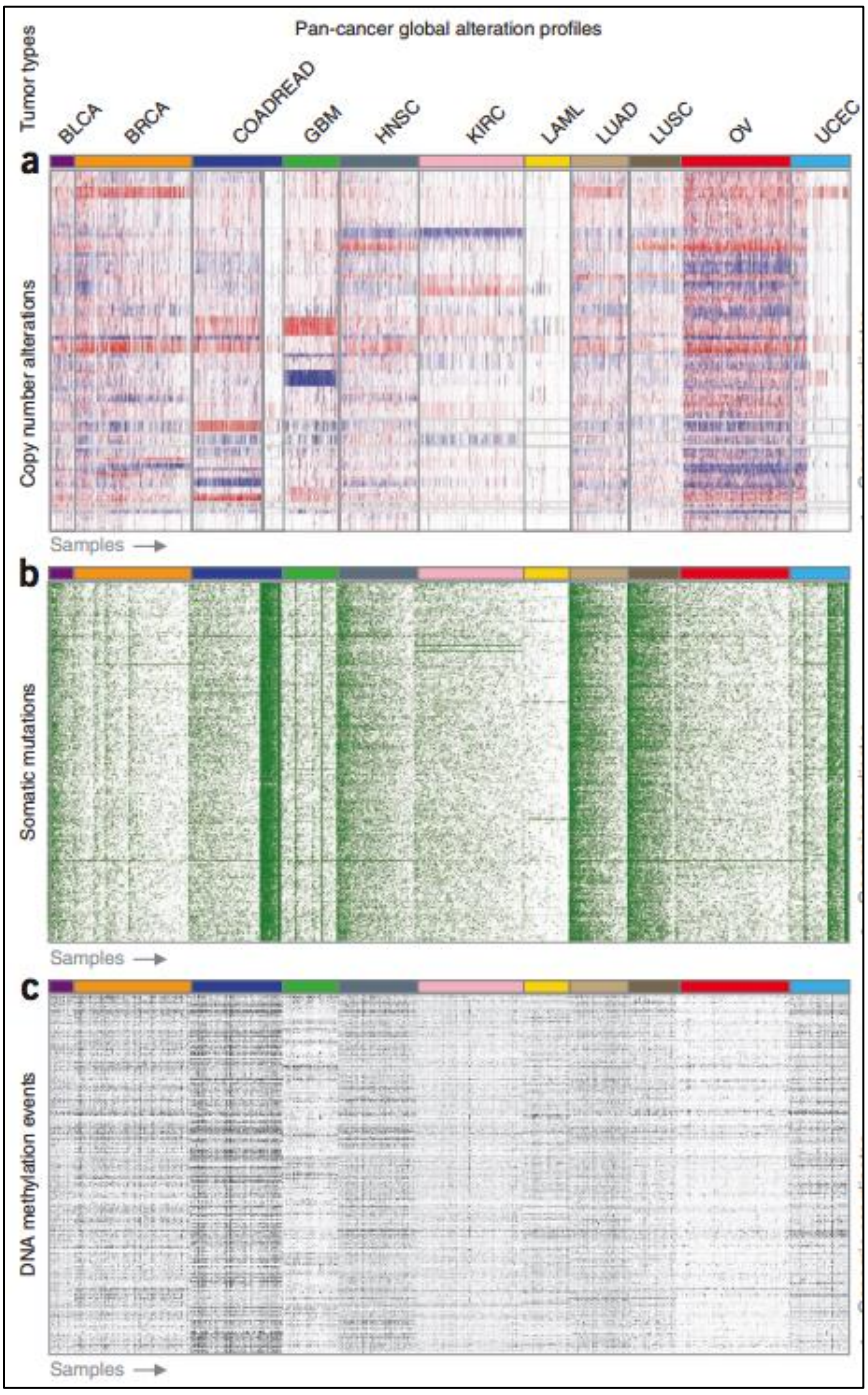


Figure 1.3 Global profile of genomic alteration of HGS-EOC (OV) in comparison with other cancer types. BLCA: bladder urothelial carcinoma, BRCA: breast invasive carcinoma, COADREAD: colon and rectum adenocarcinoma, GBM: glioblastoma multiformae, HNSC, head and neck squamous cell carcinoma, KIRC: kidney renal clear-cells carcinoma, LAML: acute myeloid leukemia, LUAD: lung adenocarcinoma, LUSC: lung squamous cell adenocarcinoma, UCEC: uterine corpus endometrioid carcinoma). A: Copy Number Alternation (CNA) profile. B: Somatic mutation profile C: DNA methylation events (22).

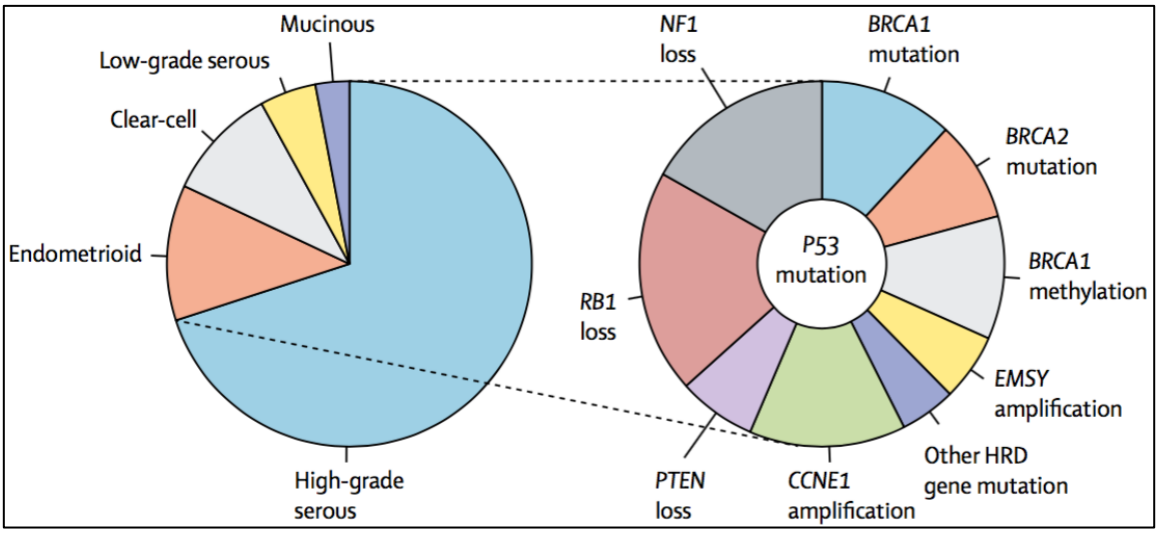


Figure 1.4 Common molecular abnormalities in HGS-EOC (9)

About ten years ago HGS-EOC was classified on the basis of gene expression profile in four main subtypes (C1-mesenchymal, C2-immune- C4-differentiated and C5-proliferative) (24). However, recently another type of HGS-EOC classification was proposed on the basis of copy number signature. In particular, Macintyre and colleagues subdivided HGS-EOC in seven different subclasses, considering six fundamental copy number features (break point count for 10 Mb, copy number, copy number change point, breakpoint count for chromosome arm, length of chains of oscillating copy number, segment size). These seven copy number signatures are associated with different mutational profile (25) and different prognosis, although a prospective validation is required to be used in a prognostic way (26).

All the studies reported above, as well as molecular features that are usually considered in clinic (i.e. *BRCA* mutations for PARPi treatment, paragraph 1.3.3) are based on the analysis of single biopsies taken in the primary tumour site at time of diagnosis. However, it is well known that HGS-EOC is characterized by high intra-tumoural, inter-tumoural, and temporal heterogeneity. Few years ago our group has in fact demonstrated that from a mutational point of view the primary tumour site and its synchronous lesions show an overlap of only 5% (27). Moreover a longitudinal analysis between matched primary tumour and tumour relapse have revealed that only 2% of mutations were conserved (28).

1.3 Clinical management of High-Grade Serous Epithelial Ovarian Cancer

1.3.1 Diagnosis

The late diagnosis that characterized HGS-EOC is due to the fact that symptoms associated to the presence of the disease are not specific, and that is the reason why often the EOC is defined as “silent killer”. Symptoms include abdominal bloating, nausea, change in bowel function, early satiety, abdominal distension, back pain, urinary symptoms, fatigue and loss of weight, and they generally appear months before EOC diagnosis. The initial investigation based on the measures of CA-125 serum protein levels and the TVS, should be associated to Computed Tomography (CT) of chest, abdomen and potentially to pelvic-abdominal Magnetic Resonance Imaging (MRI) to define the disease extension (29).

Once confirmed the EOC diagnosis, the standard EOC treatment is a multidisciplinary approach, based on the association of surgical intervention and a poly-pharmacological platinum (Pt)/taxane-based chemotherapy, to which is possible to associate the anti-angiogenic drug (Bevacizumab).

1.3.2 Surgery

Since 1980s, in the majority of cases patients initially undergo to a Primary Debulking Surgery (PDS) followed by six cycles of chemotherapy. The PDS has both a diagnostic and a therapeutic aim. In fact, on one hand allows to obtaine i) the tumour biopsies that will be analysed by the anatomo-pathologist for the histological EOC classification and ii) to determine the staging of the disease, and on the other hand to remove the tumour masses. The PDS include a variety of

surgical procedures including hysterectomy, bilateral salpingo-oophorectomy, omentectomt, lymph nodes sampling and the debulk of macroscopic tumour lesions that could lead i.e. to peritoneal resection, diaphragm stripping, partial pancreatic and hepatic resection or splenectomy. The extension of surgical procedures performed during PDS depends on the extension of the diseases, the location of tumour masses, the patients' general condition and comorbidities (30). The goal of the PDS is the complete resection of the macroscopic disease, as it is well demonstrated that the residual tumour (RT) at the end of the surgical intervention represents one of the main prognostic parameters (31). Another type of surgical approach is the administration of three cycles of chemotherapy, defined as NeoAdjuvant ChemoTherapy (NACT), followed by an interval debulking surgery (IDS) and finally other three cycles of chemotherapy. However, there are no specific clinical criteria that establish to adopt PDS rather that IDS, and the choice between the two approaches is still controversial.

1.3.3 Chemotherapy

1.3.3.1 First-line chemotherapy

Front line platinum-based chemotherapy is the mainstay for patients with diagnosis of EOC. Historically, ovarian cancer was one of the first malignancies to be successfully treated with cytotoxic chemotherapy. In particular, the alkylating agents developed in 1950s represented the first class of cytotoxic drugs used for the ovarian cancer treatment. Although many of these drugs, including melphalan, thiotepa and cyclophosphamide, demonstrated a good single-agent activity, it was promptly observed that the most effective strategy would be to employ these agents in combination, both to exploit a synergistic effect and to reduce the risk of the disease acquiring chemoresistance. Following a temporal progression, in the 1970s many combinations of drugs were in use for the treatment of ovarian cancer, with the most popular protocol

consisting of the use of cyclophosphamide and doxorubicin, in addition to methotrexate and 5-fluorouracil (32).

However, since the late 1970s the interest has headed to platinating agents for the treatment of ovarian cancer, so much so that in the last decades the standard-of-care for EOC has been referred to as "platinum-based chemotherapy". Cisplatin was the first platinum agent approved by FDA (Food and Drug Administration) in 1978 and it was tested for the EOC treatment either as a single agent or in combination with other drugs, mainly cyclophosphamide. During the 1980s another platinum compound was developed, the carboplatin (CBDCA), which demonstrated a comparable efficacy but a more favourable toxicity profile than cisplatin. As well as cisplatin, also carboplatin was tested for EOC treatment in combination with other cytotoxic agents, in particular with a new class of drugs that act as mitotic inhibitors stabilizing microtubules structures, the taxanes, of which paclitaxel (PAC) is still the prototype.

It was demonstrated by two independent randomized trials that cisplatin-paclitaxel combination had an 11% of advantage in overall-survival comparing with cisplatin-cyclophosphamide (33). Considering these results and several trials that demonstrated that CBDCA has comparable effectiveness but with a far more favourable toxicity profile than cisplatin (34) (35) (36), the standard first-line chemotherapy treatment for EOC, independently from the histotype, is represented by the association between CBDCA and PAC by intravenous administration every three weeks for 6 cycles (table 1.3):

Standard first-line chemotherapy treatment	
Carboplatin (CBDCA) [area under the curve 5-6] + Paclitaxel* (PAC) [175 mg/m^2 over 3h]	
Intravenous administration	
Once every three weeks	6 cycles

Table 1.3 Standard first-line chemotherapy treatment for EOC (32). * In cases with an increased risk of neurotoxicity PAC can be replaced with pegylated liposomal doxorubicin (Caelyx®).

More recently in addition to this standard treatment, it is possible to administer concomitantly and then as maintenance therapy an antibody directed to the Vascular Endothelial Growth Factor (VEGF), Bevacizumab, that is an FDA anti-angiogenic drug approved for ovarian cancer treatment.

In fact, two randomised trials (GOG0218 and ICON7) have suggested a benefit in terms of Progression Free Survival (PFS) (GOG0218: HR 0.9; CI 95%, 0.79 to 1.04 and ICON7: HR 0.8; CI 95%, 0.70 to 0.94), but not in terms of Overall Survival (OS) with the addition of Bevacizumab to CBDCA and paclitaxel every three weeks followed by maintenance monotherapy (37) (38).

Another class of compounds that is just recently included in clinical practice and is now under study as maintenance treatment after first-line chemotherapy is the one of poly (ADP-ribose) polymerase (PARP) inhibitors.

An international, randomized phase III clinical trial (SOLO-1) has demonstrated the efficacy of the PARP inhibitor Olaparib (Lynparza®) as a maintenance treatment in patients with newly diagnosed advanced HGS-EOC, endometrioid ovarian cancer, primary peritoneal cancer, or fallopian tube cancer with somatic or germline *BRCA1/BRCA2* mutations who had a complete o partial clinical response on the basis of RECIST 1.1 criteria (39) after first-line Pt-based chemotherapy. In particular, it was demonstrated that after a median follow-up of 41 months the risk of disease progression or death was 70% lower in patients treated with Olaparib (300 mg twice daily) than in those who received placebo (HR 0.30; CI 95%, 0.23 to 0.41) (40).

1.3.3.2 Treatments at recurrence disease

The clinical response to first-line chemotherapy is heterogeneous and not predictable at time of diagnosis. Despite more than 80% of HGS-EOC patients initially respond to first-line chemotherapy treatment, the majority of them relapse within 18 months. Actually, there are no clinical or molecular biomarkers able to predict the sensitivity to Pt-based treatment of the recurrent disease, and it is empirically based on the time lagging between the end of first-line therapy and tumour relapse (figure 1.5). On the basis of time to relapse after Pt- based treatment patients can be classified in:

I. Platinum sensitive: patients relapsing after more than 12 months from the end of chemotherapy. Generally at time of relapse those patients are re-treated with Pt-based therapy, showing a response rate to platinum re-challenge that is over than 50% (41).

II. Partially sensitive: patients relapsing within 6-12 months from the last cycle of chemotherapy. For the second line-therapy they are generally challenged again with platinum, although the response rate is lower in comparison to platinum-sensitive patients.

III. Platinum resistant: patients relapsing within 6 months from the end of therapy. They are not treated with platinum again.

IV. Platinum refractory: patients showing a stable or a progression of disease (PD) during Pt-based treatment. The PD is defined by Response Evaluation Criteria In Solid Tumour version 1.1 (RECIST 1.1) as an increase of at least 20% in the sum of

diameters of target lesions taking as reference the smallest sum on study. Also the appearance of one or more new lesions is considered progression (42).

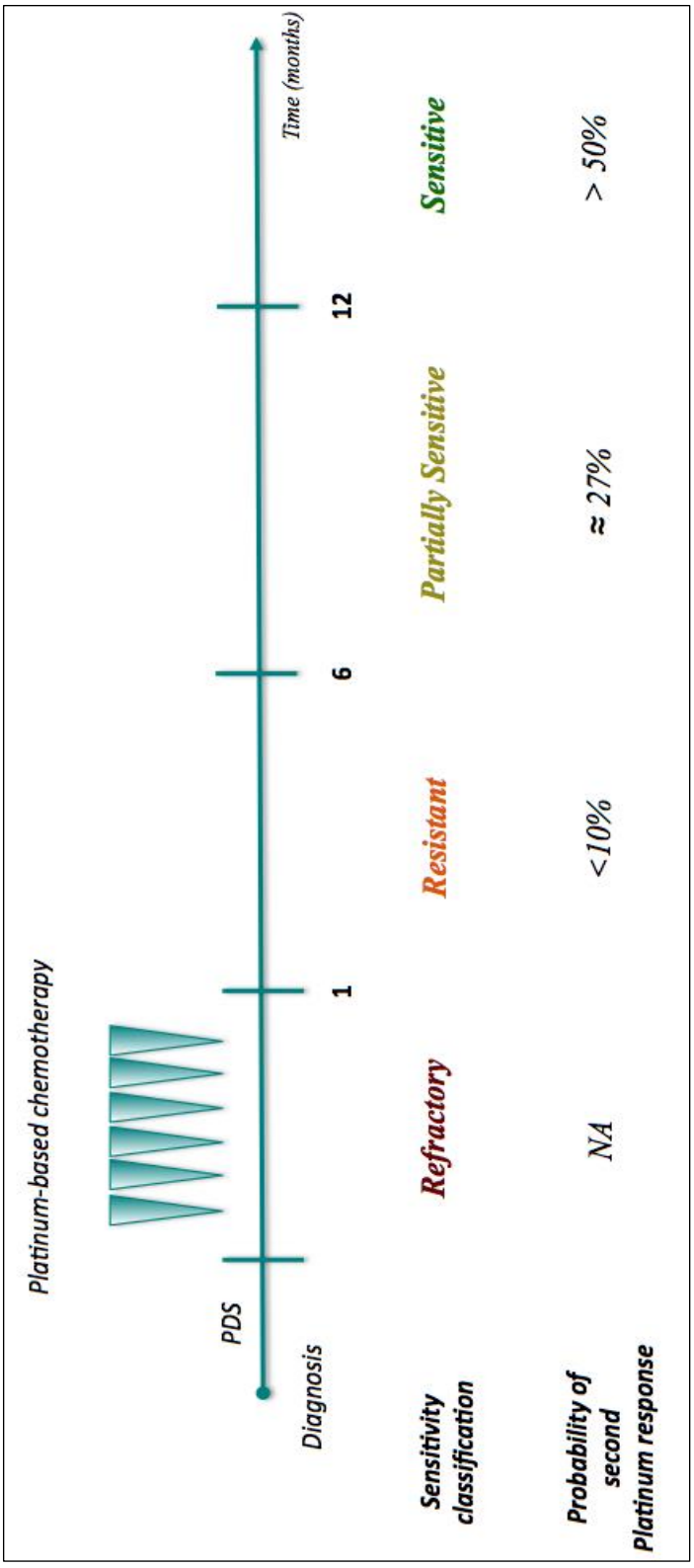


Figure 1.5 Patients' classification on the basis of their response to platinum-based chemotherapy. PDS: Primary Debulking Surgery

On the basis of this classification different therapeutic options are available at time of recurrence:

- Generally platinum sensitive patients are retreated with CBDCA-PAC chemotherapy. Alternatively to PAC it is possible to administrate in combination with CBDCA, gemcitabine or pegylated liposomal doxorubicin (PLD, Cealix®), that have demonstrated both improvement in PFS (HR 0.82, CI 95%, median PFS 11.3 vs 9.4 months) and a better toxicity profile compared with the standard treatment CBDCA-PAC (43). Moreover, in addition to Pt-based chemotherapies described above, in platinum sensitive recurrence is possible to administer other two classes of drugs: anti-angiogenic agents and PARPi. Several clinical trials had demonstrated the efficacy in terms of PFS combining CBDCA-PAC (GOG213 trial) and CBDCA-gemcitabine (OCEANs trial) with bevacizumab followed by maintenance monotherapy (44) (45). Results of MITO16B-MaNGO-OV2B-ENGOT OV17 trial, presented at ASCO 2018, demonstrated that the re-challenge with bevacizumab in combination with platinum-based doublets is associated with a significantly prolonged PFS also in patients treated with bevacizumab in the first-line/maintenance chemotherapy.

  Regarding the use of PARPi, results reported in literature derived from mainly two clinical trials, indicate that, although the major benefit derived from their administration are obtained in *BRCA* mutated and Homologous-recombination deficient (HRD) positive patients, as expected (46), improvement in PFS is also observed in BRCA wt and HRD-negative cases. To these reasons, niraparib (Zejula®) and rucaparib (Rubraca®) have been approved by FDA and European Medical Agency (EMA) as maintenance therapy of platinum sensitive ovarian cancer recurrence in response to platinum-based treatment, independently from *BRCA* genes and HRD status (47) (ARIAL3 trial, (44)). Also, olaparib (Lynparza®) has been approved as maintenance therapy for adult patients with recurrent HGS-EOC who are in complete or partial

response to Pt-based treatment.

- Platinum partially sensitive patients with disease recurrence, can be treated using the same pharmacological settings available for platinum sensitive cases, included the maintenance treatments with bevacizumab, rucaparib, niraparib or olaparib.

  The combination of trabectedin (ET-743, Yondelis®) with PLD is considered a valid therapeutic FDA/EMA-approved approach for relapsed platinum partially sensitive patients. The INOVATYON phase III trial is aimed at demonstrating whether there is an OS advantage by the sequential use of trabectedin/PLD followed, at relapse, by platinum re-challenge, over carboplatin/PLD combination. The primary results analysis of INOVATYON trial are expected in August 2020.

- For platinum refractory and resistant patients the therapeutic approaches are based on single agent-treatment. The chemotherapeutic agents usually administered in these cases are PLD, gemcitabine, taxol (high-dose density, weekly administration) or topotecan. The AURELIA trial has demonstrated a benefit in terms of PFS and quality of life with the association of PLD, gemcitabine or taxol with bevacizumab followed by maintenance therapy until progression of the disease, for patients which have never been received bevacizumab before (AURELIA) (48). This association is approved by EMA.

1.4 The issue of drug resistance in High-Grade Serous Epithelial Ovarian Cancer

The main problem related to HGS-EOC is the resistance to Pt-compounds that sooner or later patients develop.

In fact, although about 80% of patients are initially sensitive to the first-line Pt-based therapy

the majority of them relapse with a recurrent disease that usually follows a frequent relapse-response pattern before becoming definitely resistant to Pt-based treatment. Although our knowledge about the mechanisms underlying the therapy resistance in EOC is increased over the past decades, we are only beginning to appreciate the complexity of this issue. The complexity is both clinical, as there are no biomarkers able to predict the response to Pt-based treatment and biological, due to the various mechanisms involved and to the marked molecular heterogeneity that characterized HGS-EOC.

Since the most important drugs for treatment of HGS-EOC are platinum compounds, in this thesis the attention will be focused on the resistance mechanisms related to this class of drugs.

Considering the timing of Pt-resistance development it is possible to distinguish two different types of resistance (figure 1.6):

i) intrinsic (or primary)

ii) acquired resistance

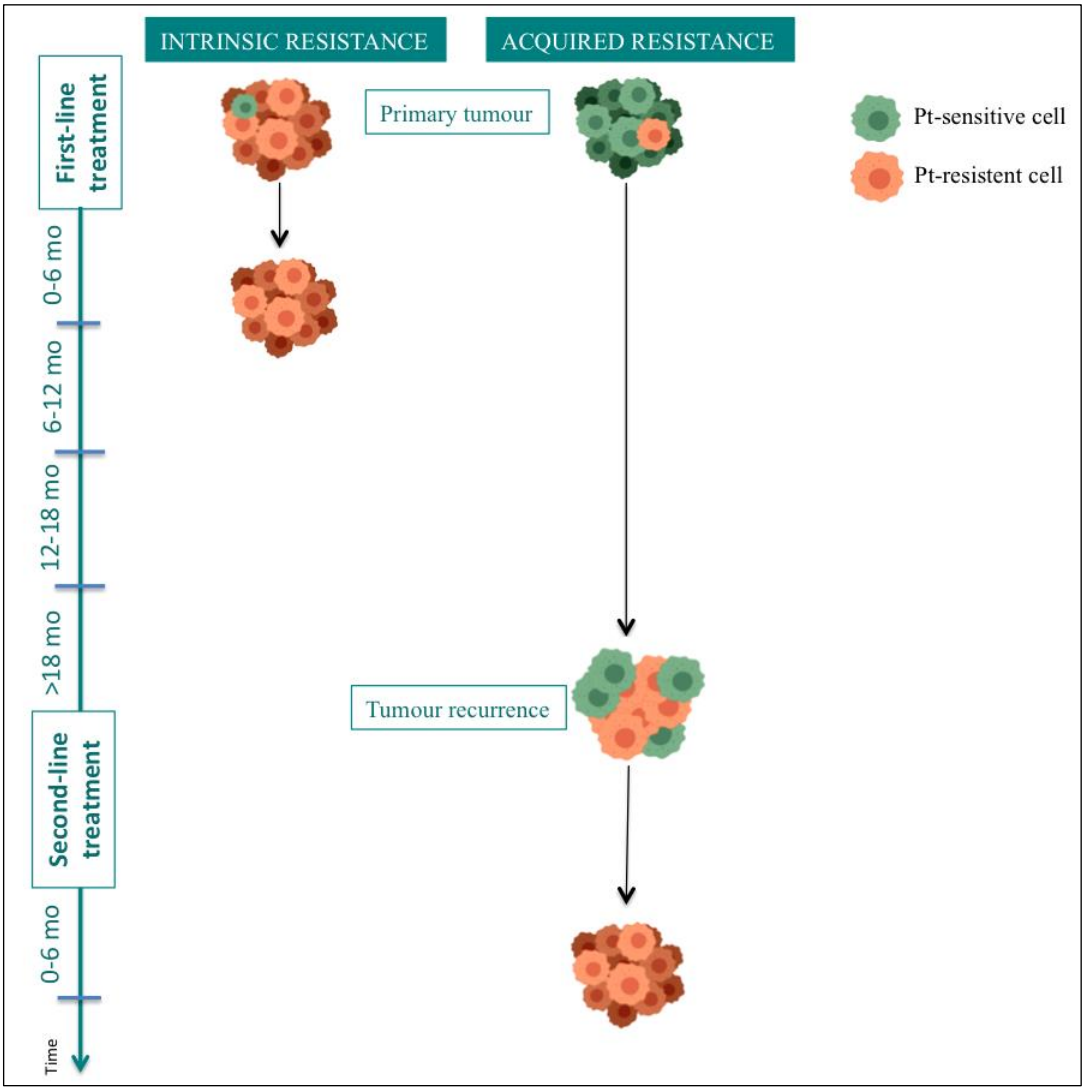


Figure 1.6 Intrinsic and acquired Pt-resistance. Mo, months

The primary resistance to Pt-compounds, involving about 15-25% of HGS-EOC patients, is referred to those tumours that do not respond or relapse within 6 months from the end of the first-line platinum-based treatment (paragraph 1.3.3.2). The primary resistance, is also named "intrinsic" resistance, as the biological features related to non-response to therapy are already present in some cells (intratumoural heterogeneity) composing the tumour masses at time of diagnosis (49).

In contrast, the acquired platinum resistance has been defined as the failure to respond to subsequent treatment having previously demonstrated sensitivity to one or more lines of pt-based chemotherapy.

The multiple biological mechanisms involved in platinum resistance will be described in detail

in paragraph 1.4.2, after a brief description regarding the mechanism of action of platinum drugs (paragraph 1.4.1).

1.4.1 Mechanism of action of platinum compounds

It was in the 1960s that Barnett Rosenberg (figure 1.7) at the Michigan State University accidentally discovered cisplatin (CDDP), the first platinum-derivate compound approved for clinical use. While he was studying the effects of electric dipole fields on cell division, he noticed that a product released by platinum electrodes due to electrolysis reaction, was able to inhibit DNA synthesis in microorganisms.

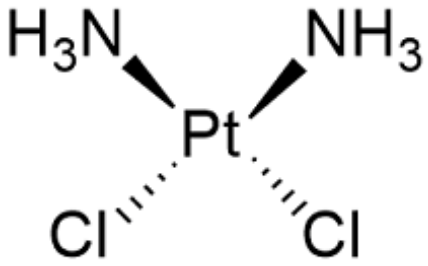


Figure 1.7 Barnett Rosenberg (1926-2009) and cisplatin structure

The following detailed chemical analysis purified and identified the compound, named as cisplatin (*cis*-[Pt(II)(NH(3))(2)Cl(2)] ([$PtCl_2(NH_3)_2$] or CDDP) (50). Cisplatin is a metallic (platinum) coordinator compound with a square-planar geometry. It contains two chloride leaving groups in *cis* dispositions, that remain bound to the platinum in the plasma due to its high chloride concentration (about 100 mM). Once inside the cell, where the chloride concentration sharply decreases (4-12 mM), the cisplatin undergoes aquation of one or two of

the two chloride leaving groups (figure 1.8), leading to the formation of activated complexes $[Pt(NH_3)_2Cl(H_2O)]^+$ and $[Pt(NH_3)_2(H_2O)_2]^{2+}$. To note the chemical reaction that generates the activated aqua platinum species is the same for cisplatin and for both carboplatin and oxaliplatin, the only two FDA approved cisplatin-derivatives.

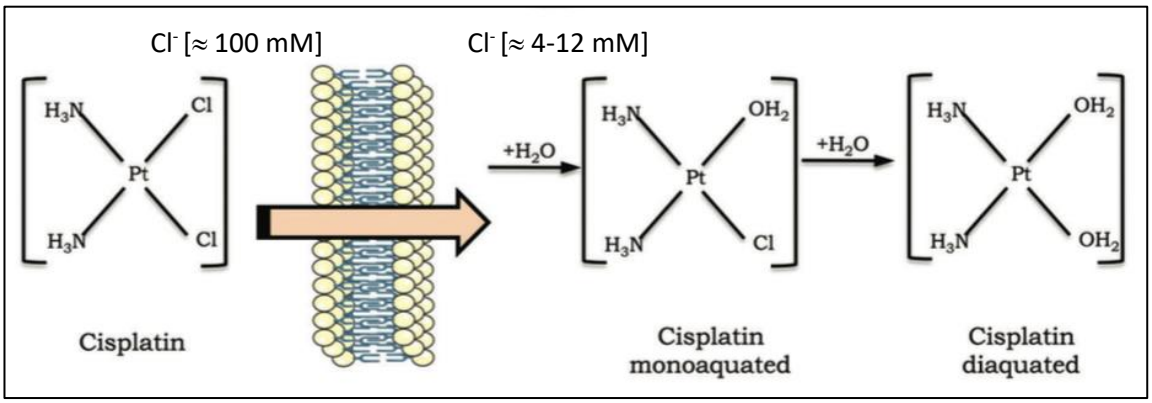


Figure 1.8 Cisplatin activation process (51).

These activated complexes are able to bind RNA, DNA, proteins and membrane phospholipids, although they preferentially bind the genomic DNA inside the cellular nucleus. In particular, the most common binding site is represented by the N7 atom of purine bases, with a very predominant preference of guanine over adenine. Although the chemical structure of activated complexes allows monofunctional bindings to a single purine base, or DNA-protein cross-link, the anticancer activity of platinum-compounds is related to the formation of *intra-* (binding of two guanines located on the same DNA strand) and *inter-* (binding of two guanines located on opposite DNA strand strands) covalent bis-adducts (figure 1.9) (52). Although inter-strand crosslinks are formed less frequently than intra-strand crosslink, damages induced are more toxic for cells. In fact, as inter-strand crosslinks involve both complementary strands, they result more inhibitory to DNA replication and transcription compared to the intra-strand crosslinks (53) (54).

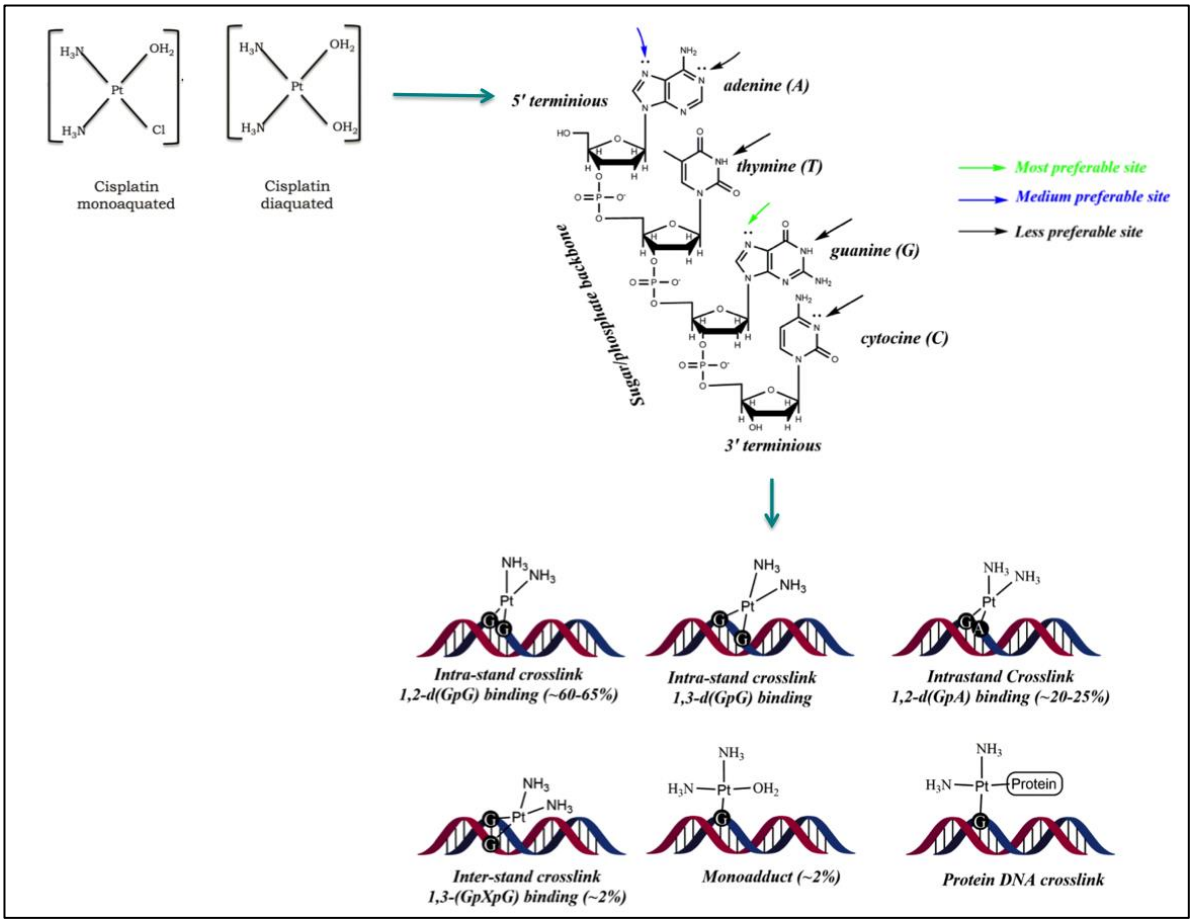


Figure 1.9 Binding between platinum-derivatives activated complex and DNA (modified from (55))

The formation of Pt-induced DNA adducts leads to a block of DNA synthesis and transcription, which in turn triggers to an intricate intracellular signal transduction cascade aimed to eliminate the lesions. While intra-strand cross-links cause a distortion of double helix and activate mainly the nucleotide excision repair pathway (NER), inter-strand cross-links rely upon the interaction among different DNA repair pathways including NER, the Fanconi Anemia (FA) and the Homologous Recombination (HR) pathway, which is involved in the repair of DNA double strand (dsDNA) breaks induces by platinum compounds (56). In order to provide adequate time to DNA repair mechanisms to remove the lesions, cell cycle is arrest, but in case of impaired repair or serious damage the cells undergo apoptosis (figure 1.10).

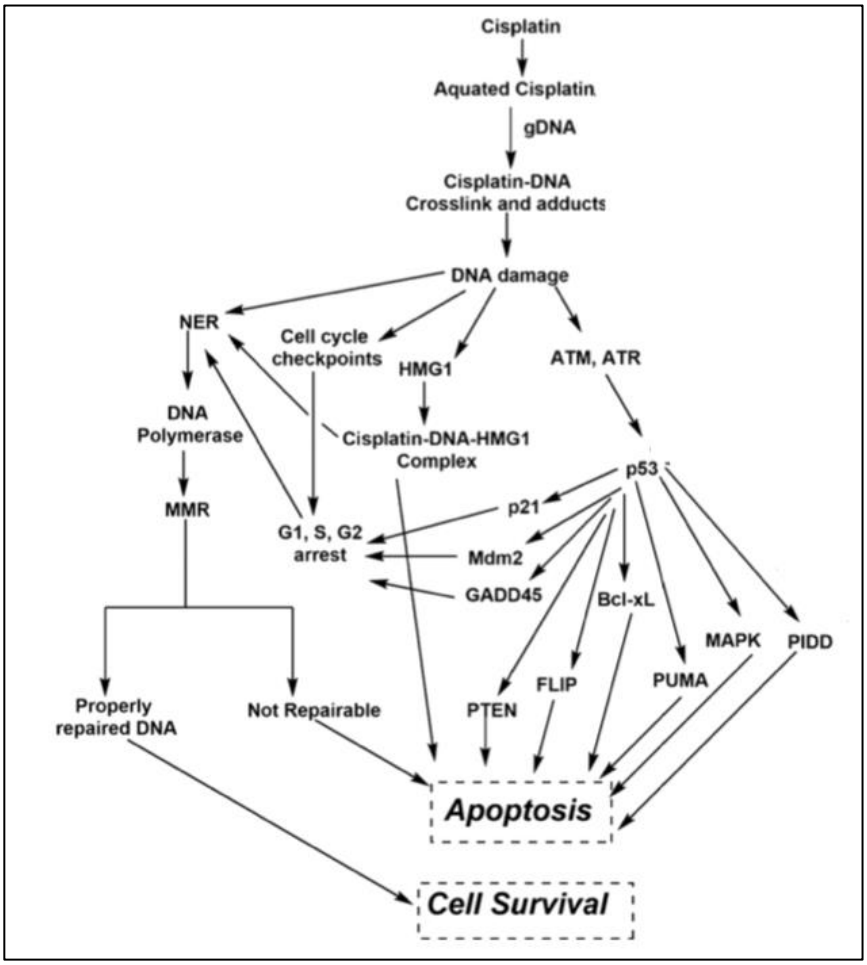


Fig.1.10 Mechanism of action of platinum-compounds (modified from (55))

1.4.2 Mechanisms of resistance to platinum compounds

The issue of primary and acquired resistance to platinum compounds is extremely intricate and involved many biological processes.

In this section the major biological mechanisms involved in Pt-resistance are cited, but a particular attention will be paid to the role of *BRCA1/BRCA2* and the HR pathway. The mechanisms are classified on the basis of alterations that: i) involve the steps upstream the platinum-DNA bound (pre-target resistance); ii) are directly related to the formation and recognition of Pt-DNA adducts (on-target resistance); iii) concern the lethal signalling pathways elicited by Pt-mediated DNA damages (post-target resistance).

1.4.2.1 Mechanisms of pre-target resistance

There are three fundamental mechanisms by which the cells can elude the potential cytotoxic activity of platinum compound (figure 1.11), inhibiting the interaction between the drug and its main target (DNA):

i) reduction of intracellular accumulation of Pt

ii) increase of Pt efflux

iii) increase of Pt intracellular sequestration

Although for long time it has been believed that cisplatin entered into cytoplasm cells' through passive diffusion, in the last decades transmembrane proteins involved in copper homeostasis have emerged to play a very significant role in cisplatin uptake and its consequent intracellular accumulation (57). In particular low expression levels and polymorphisms (58) of the copper transporter 1 (CTR1), which is mainly responsible of cisplatin uptake, are associated with resistance to Pt-based treatment both *in vivo* and *in vitro* (59) and with a poor OS and PFS (60). Moreover the exposure to clinical relevant concentration of cisplatin has been reported to lead to CTR1 internalization and degradation, contributing at least in part for multiple instances of acquired resistance (61).

The resistance to DDP is also be related to increase drug export. There are two main classes of transporters involved in cisplatin cellular export: the ATP binding cassette (ABC) transporters and P-type ATPase transporters. Multi resistance associated proteins (MRP), belonging to ABC transporter are transmembrane proteins responsible for the efflux of glutathione-platinum conjugates. In particular, many years ago, several studies had pointed out the role of MRP2 (also known as cMOAT, canalicular multispecific organic anionic transporter) as the major transmembrane protein responsible for DDP efflux (62). Although the association between overexpression of MRP2 and resistance to platinum-based chemotherapy has been

demonstrated years ago in several solid tumours i.e. colorectal, oesophageal, hepatocellular (63) (64) (65), only recently it has been confirmed also in ovarian cancer (66). As previously anticipated another class of transmembrane proteins are involved in cisplatin efflux that is the copper (Cu) extruding P-type ATPase transporters class, in particular ATP7A and ATP7B transporters. Their up-regulation is associated to resistant phenotype to cisplatin in ovarian cancer cell lines and it was demonstrated to be associated to prognosis in ovarian cancer patients (67). The altered activity and expression of ATP7A and ATP7B transporter seems to be related with the intracellular copper concentration, indicating that intracellular Cu homeostasis would influence the efficacy of platinum-based drugs (68).

Finally, the last pre-target mechanism of resistance to DDP regards the cytoplasmic conjugation of aquated cisplatin with nucleophilic species including glutathione (GSH) and the binding with metallothioneins (MT) proteins. Both in vivo and in vitro models, have demonstrated that elevated levels of GSH, of the enzymes that are involved in GSH synthesis (γ-glutamycysteine synthetase) or of enzymes that coniugate the aquated-DDP and GSH (glutathione S-transferase), are associated to DDP resistance. The conjugation of aquated-DDP with GSH, in fact, leads to the cellular export by MRP transporters (69). Also the metallothioneins, in particular MTI and MTII, are involved platinum resistance, due to their capabilities to act as cytoplasmic DDP scavengers. It has demonstrated that their overexpression is related to platinum resistance and poor prognosis in ovarian cancer (70).

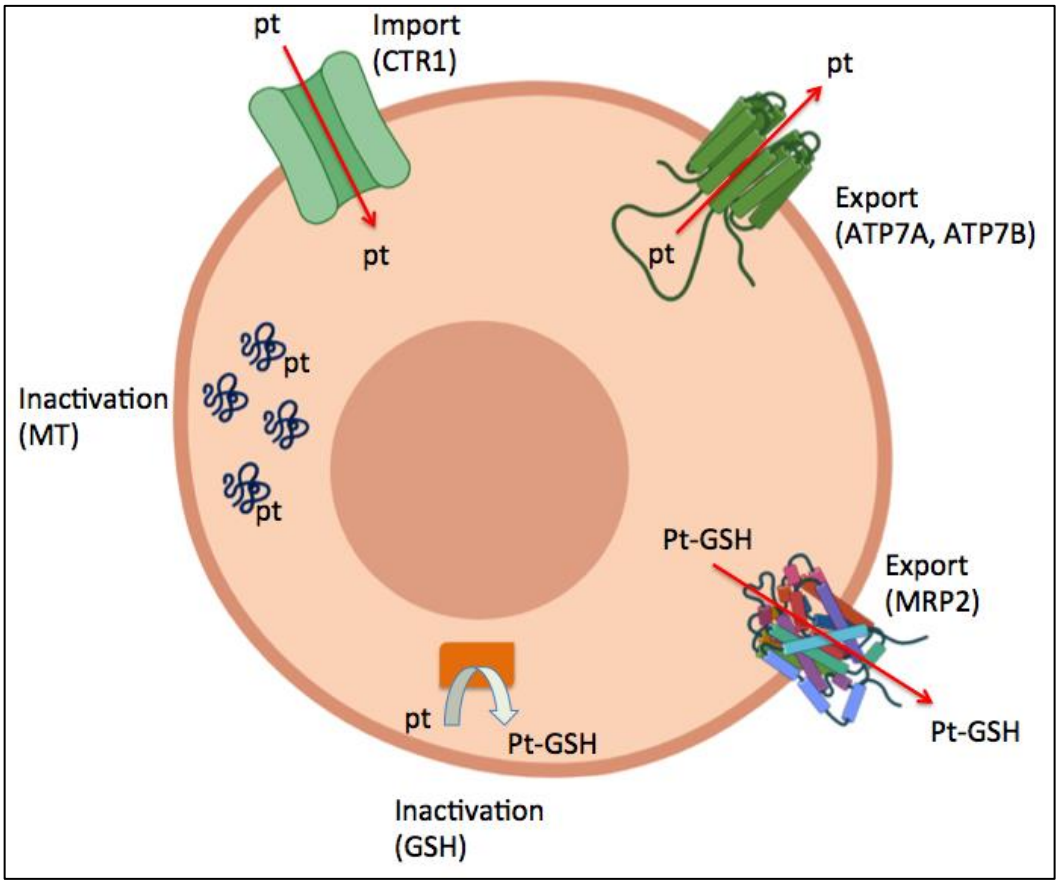


Figure 1.11 Mechanisms of pre-target platinum resistance

1.4.2.2 Mechanisms of on-target resistance

Once overcome the pre-target resistance mechanisms described above, DDP is able to carry out its cytotoxic activity by binding DNA and forming Pt-DNA adducts. As DNA represents the main target of Pt -based drugs, the sensitivity/resistance to therapies is strictly related to the capability of the cells to recognize and repair the platinum-induced DNA damages. For these reasons, alterations (mutations, copy number or epigenetic changes) of genes involved in DNA repair mechanisms are associated with increased or reduced drug sensitivity.

The on-target resistance in HGS-EOC is mainly due to:

- increased nucleotide excision repair (NER) proficiency
- mismatch repair (MMR) deficiency
- increased HR proficiency

As alterations in NER, MMR involve small percentage of HGS- EOC (8% and 3% respectively) (71) and nowadays no drugs are available to target or exploit defects in these pathways, they will be very briefly described. On the contrary, a particular attention will be paid to defects in HR pathway, which involved about 50% of HGS-EOC patients and whose alterations are exploited by a new class of drugs (PARPi) recently approved for HGS-EOC treatment (see paragraph 1.3.3).

The majority of DDP-induced lesions, both intra and inter-strand, are removed by NER, which is the main DNA repair pathway involved in the resolution of "bulky lesions". In this setting, after the formation of platinum-induced adducts, the damaged nucleotides are excised from DNA upon cleavage of both sides of the lesion (XPG-3' and XPF-ERCC1-5' nuclease), followed by DNA synthesis.

Early reports correlate an increased in expression levels of mRNA *ERCC1,* with the platinum resistance in ovarian cancer cell lines and tumour tissue (72). However, conflicting results published in literature regarding the expression of ERCC1 and platinum response are most likely due to the facts that 1) many studies did not investigate the specific *ERCC1* isoform (ERCC1-202) involved in the formation of active XPF-ERCC1 complex (73) 2) the evaluation of expression levels of a single protein involved in a complex DNA repair system such as NER could not totally reflect the status of entire pathway. Further detailed pre-clinical and clinical investigations are still needed.

In contrast to NER, defects in MMR were associated to resistant phenotype, for the ability of MMR-related proteins, in particular *MSH2* and *MLH1* often mutated or underexpressed, to recognised the DNA damages and activate apoptotic processes (74).

However also in this case results reported in literature are discordant in associating MMR deficiency with Pt-resistance in HGS-EOC, and the role of MMR inactivation in treatment response deserves further investigation (75).

To resolve the formation of inter-strand crosslinks and the double strand break (DSB) DDP-induced, another important mechanism of DNA repair is involved: the homologous recombination (HR) pathway.

HR is an error-free pathway, strictly activated in phase S-G2, that requires the presence of sister chromatid as a template to solve the double strand breaks (DSB) (figure 1.12). In particular, once that DSB has been recognized by MRE11A-NBS1-RAD50 (MRN complex), ATM recruitment and ATM-mediated phosphorylation lead to MRN activation. This event causes the resection of 5' strand-ends on either sides of DSB determining, via BRCA1-dependent processes, the exposure of the two 3' single-strand DNA (ssDNA) regions. Through the action of BRCA2 and PALB2 proteins, RAD51 is able to bind the exposed ssDNA 3'end and to invade the double helix of the intact homologous stretch of DNA on the sister chromatid. DNA polymerases use the 3'-ssDNA end as a primer for the synthesis of the new DNA strand, using the homologous DNA sequence as a template with the formation of a displacement loop (D-loop) (76). Now the second DSB end can be captured forming an intermediate structure with two Holliday junctions (HJ). The structure is now solved by DNA Polymerases and DNA ligases at the HJ level through a crossover or non-crossover mode, that means with or without the exchange of genetic material between two homologous chromosomes non-sister chromatids. Alternatively, it can proceed to the annealing of the ssDNA-end portion, followed extension and ligation phases. In this case the final product is always not crossover (77).

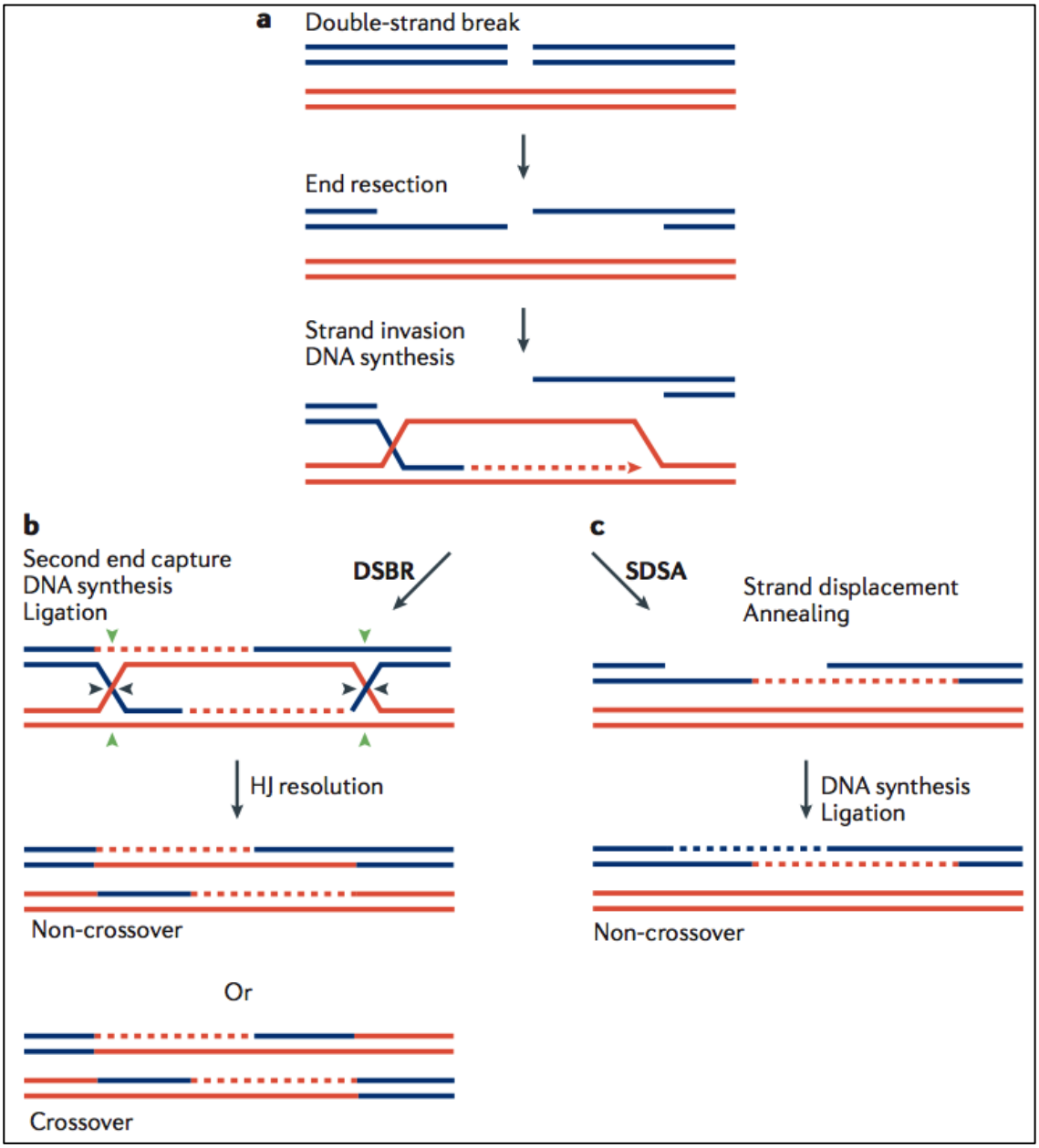


Figure 1.12 Mechanism of action of Homologous Recombination (HR) pathway to repair double-strand break DNA (77).

As previously anticipated, HR is involved in the resolution of inter-strand adducts that lead to DSB, so as expected an increase in HR activity allows cells to repair DNA damages induced by DDP, thus reducing DDP cytotoxic activity. On contrary defects in HR make the tumour cells unable to repair Pt-induced DNA damage, leading to an increase genomic instability, that causes activation of apoptosis pathways or mitotic catastrophe (78).

As reported in figure 1.13, about 50% of HGS-EOC cases are characterized by defects in HR pathway. In particular, two key genes for HR system, that are *BRCA1* and *BRCA2* are mainly

involved and account for the 30% of all HR deficient cases. *BRCA1* and *BRCA2* are affected by germline mutations, associated to Breast-Ovarian cancer syndrome (paragraph 1.1.1) in 8% and 6% of HGS-EOC cases, whereas somatic mutations are detected in about 3% of cases for both genes (19). The majority of these mutations are frameshift or indel (insertion or deletion) and are frequently associated to heterozygous loss (LOH) (81% in *BRCA1* and 72% in *BRCA2*), indicating that both alleles are inactivated, so the function of the proteins is completely lost (79). Mutations in *BRCA1* and *BRCA2* are mutually exclusive with *BRCA1* epigenetic silencing, that occurs via promoter hypermethylation (about 10% of cases). Other somatic mutations that affect the proficiency of HR include mutations in *CDK12* (3% of cases), which lead to HR deficiency through transcriptional suppression of several HR genes included *BRCA1*, *RAD 50, RAD51, RAD51C, RAD54L*. Other mutations were observed in DNA damage response genes involved in HR (*ATM, ATR, CHEK1* and *CHEK2*) and Fanconi Anemia (FA) genes (mainly *PALB2, FANCA, FANCI, FANCL* and *FANCC*).

Moreover, alterations that affect genes that indirectly regulate HR activity are reported. For example focal deletion in genomic region (10q23.31) that include *PTEN* gene, which is reported to be associated with a downregulation of RAD51, has found in about 7% of HGS-EOC cases (79).

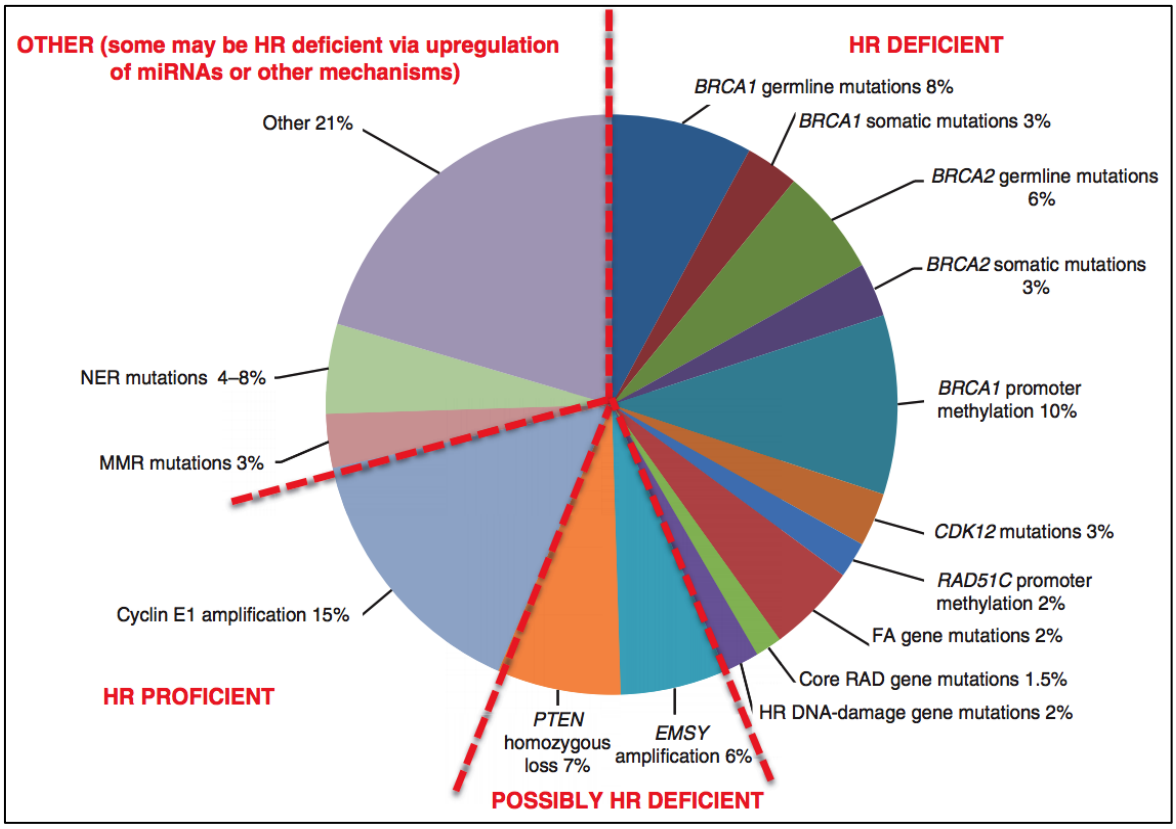
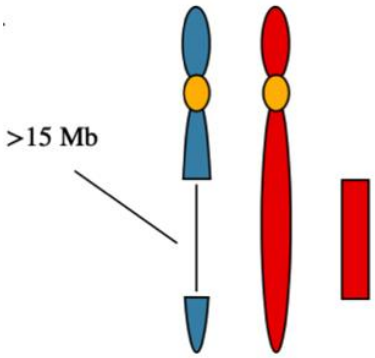


Figure 1.13 Genetic and epigenetic alterations in DNA repair pathways in HGS-EOC.

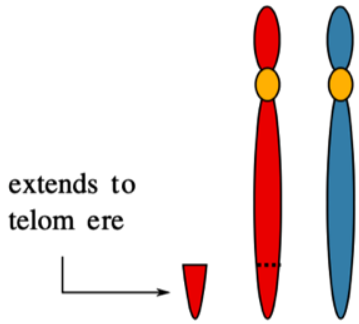
From data published in literature it was well established that HGS-EOC patients with somatic or germline mutations in *BRCA1* and *BRCA2* genes are associated with increased sensitivity to platinum compound and PARPi and with an improved survival outcome (80) (46). Pathogenic mutations in *BRCA1* and *BRCA2* are most commonly associated with the Homologous Recombination Deficiency (HRD), however alterations in other HR-related genes can cause the dysfunction of the entire pathway. In the last years the increasing interest in PARPi, has led to considere not only the mutational status of BRCA1/2 but also the entire status of HR DNA repair pathway. Although initially a gene expression signature associated to HR status was discovered analyzing by microarray approach on HR-proficient and HR-deficient cell lines (81) the development of next generation sequencing technology allowed to study more in details the HR status, analyzing genomic alterations (genomic "scars") that represent a consequence of increase in genomic instability due to HR dysfunction.

The three scars considered are:

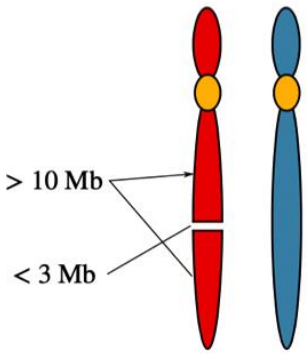
- loss of heterozygosis (LOH) in genomic region >15 Mb, but that not involve the entire chromosome (82)



- telomeric allelic imbalances (TAI) of at least 1 Mb



- large-scale transition (LST) that is a chromosomal break between adjacent regions of at least 10 Mb, with a distance between them no longer that 3 Mb



On the basis of these three genomic features, it is possible to calculate an HDR score (figure 1.14), defined as the unweight sum of the number of LOH+TAI+LST. This approach is used by FDA-approved Myriad's test (myChoice®) to select patients eligible for PARPi treatment

(HRD score < 42= HRD negative, HRD score > 42= HRD positive).

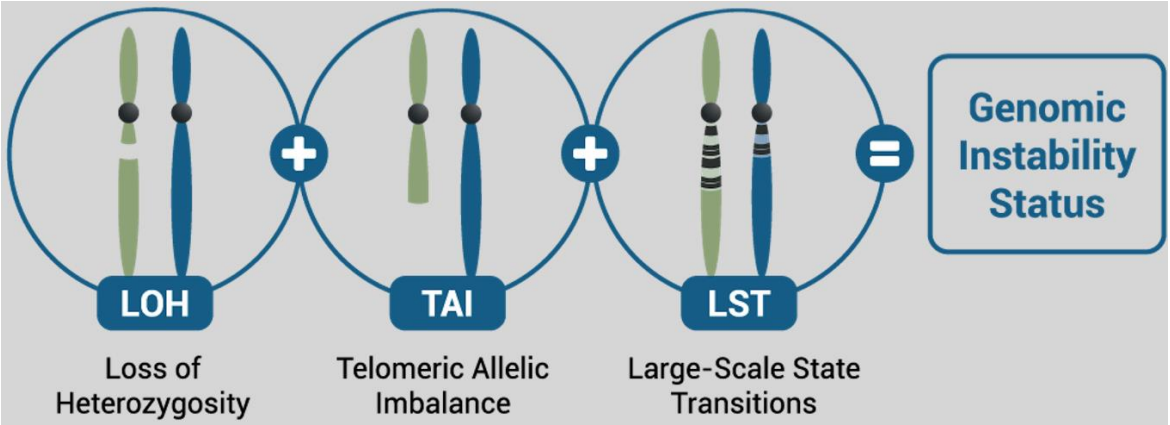


Figure 1.14 Calculation of Homologous Recombination Deficiency score

Up to now, it has been described the association between deficiency in homologous recombination pathway and the increased sensitivity to pt-compounds and PARPi, however resistance can develop through restoration of HR DNA repair system's functionality. In fact, it was demonstrated that in germline mutation carrier patients the acquisition of secondary somatic mutations in BRCA1/2 genes, which are able to restore the open reading frame and so the protein function (revers mutations), is a common mechanism to shift from HRD to HR proficient phenotype (about 33% of acquired resistance cases) (83) (84). These reverse mutations, as well as the decrease in promoter *BRCA1* methylation (85), probably due to a selective pressure related to prolonged platinum exposure, are mechanisms that can explain the acquired resistance to Pt-compounds in HGS-EOC patients.

1.4.2.3 Mechanisms of post-target resistance

In this chapter will be shortly addressed some molecular features that allow cells to avoid cellular death as a consequence of the formation of Pt-DNA adducts in HGS-EOC. Mainly, post-target resistance to Pt-compounds regards general alterations in mechanisms involved in the apoptotic pathway of signal transduction that is activated in response to DNA damage as well as in problems with the cell death executioner machinery itself. These post-target mechanisms of resistance are common to many cytotoxic agents and are based on a *disequilibrium* between pro- and anti-apoptotic cellular factors, which allows cells to avoid cellular death by increasing its genomic instability.

For example it was demonstrated in pre-clinical and clinical studies that the overexpression of several anti-apoptotic protein such as Bcl-2, Bcl-X_L, and MLC-1 correlate with platinum resistance and tumour recurrence (86) (87) (88). Also the overexpression of proteins belonging to the inhibiting apoptotic proteins family (IAP), such as cIAP1 and XIAP, has been demonstrated in preclinical studies to suppress apoptosis, supporting the chemoresistance in HGS-EOC (89).

The majority of the studies focusing on alterations in apoptotic pathways as post-target mechanism of platinum resistance, have been conducted on HGS-EOC cell lines, some of which appear to be misclassified or contaminated and that certainly represent useful but limited pre-clinical model that are not able to fully capture the complexity and the heterogeneity that characterized HGS-EOC. For this reason, it is getting clear that the translational research focusing on study directly tumour biopsies derived from patients and correlate biological features with prognostic and predictive parameters, represents a more realistic approach to investigate such a complex disease.

1.4.3 Transcriptional signatures associated to platinum response

For many years the issue of resistance to Pt-based compounds have stimulated the research of molecular biomarkers that would help to stratify patients' risk of relapse at time of diagnosis, and thus avoiding unnecessary treatments to those patients who are intrinsically resistant (see figure 1.6).

As a general comment, overall data reported till now in literature showed that a single gene approach failed to identify the few key regulatory genes responsible for therapy resistance, and this is mainly due to the complexity of regulatory mechanisms underlying therapy resistance and the high molecular heterogeneity of HGS-EOC (see paragraph 1.2). As a consequence, the scientific community was forced to move from a single gene to a pathway-level analysis.

Using integrated multi -omics approaches, Patch and colleagues, evidenced differences in the genomic profile of Pt-sensitive (Pt-s) and Pt-resistant (Pt-r) cases. In particular, as previously anticipated by Tothill (24), it was confirmed that the presence of the molecular signature C1/mesenchymal (see paragraph 1.2) was associated with both primary and acquired resistance, early relapse and poor OS (90).

Integrating genes and miRNAs expression analysis from four independent datasets, Yang and colleagues, reduced the Tothill's molecular sub-classification to two different groups associated with prognosis: the integrated mesenchymal subtype (iM) and the integrated epithelial subtype (iE). The iM subtype, significantly characterized by a shorter OS in comparison with the iE subtype , was biological defined by eight key regulatory miRNAs, including miR-200 and miR-141, which were directly or indirectly involved in the Epithelial to Mesenchymal Transition (EMT) process (91).

The EMT is a complex biological pathway involved in tissue embryogenesis that have increased importance in tumour biology over the last years. Although is beyond the scope of this thesis a detailed discussion of the role of the EMT in tumour progression towards malignancy, some

features will be briefly reviewed with a special focus on EMT's effects on platinum drug resistance.

The EMT process is characterized by a loss of epithelial cells morphology and cytoskeletal reorganization with acquisition of mesenchymal features, one of which is the ability of tissue invasion and an increased metastatic potential (figure 1.15).

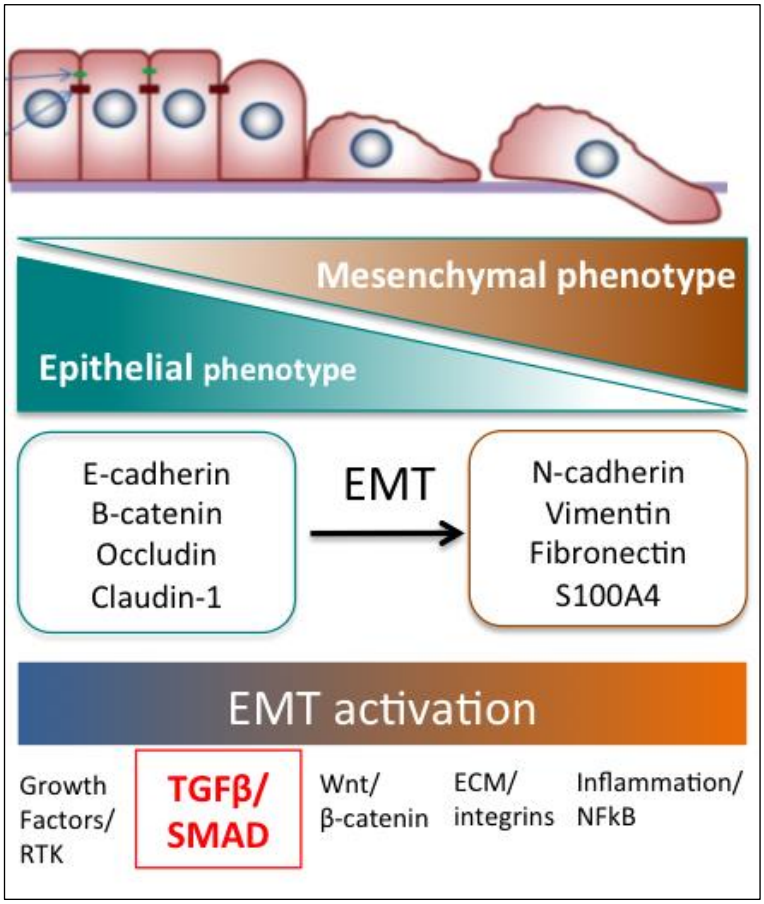


Figure 1.15 Epithelial to Mesenchymal Transition (EMT) process. In figure are reported the main markers of epithelial and mesenchymal phenotype, and the pathways involved in the EMT activation

The role of EMT in HGS-EOC Pt-resistance has long been discussed in literature. Gene expression data generated from our laboratory using 23 matched patients' biopsies, who were sensitive to Pt at time of primary surgery and became resistant after several pt-lines therapy demonstrated and validated the presence of a Pt-resistance signature indicative of EMT activation, induced by the transforming growth factor β (TGFβ) (92). In particular, as further demonstrated by Parikh et al., the activation of TGFβ-mediated EMT in HGS-EOC, occurs

through *miR-181a-5p* activity, by inhibiting its functional target SMAD-7, a regulatory inhibitor element of pathway, leads to increase in cellular survival, migration, invasion and drug resistance. In fact, high levels of *miR-181a-5p* and phosphorylated SMAD2 protein (SMAD2-p) which represent a key element of EMT TGFβ-mediated cascade, were associated to shorter time of recurrence and poor outcome in HGS-EOC patients (93). These results were confirmed by a more recent work published by our group in collaboration with the clinicians of Policlinico Gemelli in Rome (Italy), that demonstrated the prognostic and the predictive role of *miR-181a-5p* and SMAD2-p in HGS-EOC patients underwent NACT (94).

It is plausible that due to the complex and heterogeneous nature of Pt resistance, EMT pathway is just one of the molecular pathways that can explain the issue of drug resistance in HGS-EOC. In the last years research focused on the role of miRNA, as general controller of different molecular pathways in the cell. The miRNAs expression analysis conducted on HGS-EOC biopsies taken at time of primary surgery, identified a molecular signature (MIROvaR) that is predictive of risk of progression or relapse. In particular this signature, composed by 35 miRNAs, which does not include *miR-181a-5p* but involved most members of the miR-200 family that are key regulators of EMT process, is able to classify patients in "high" and "low" risk of recurrence (95).

Many other studies were conducted analyzing genome methylation profile (96) or copy number expression signatures (97) however nowadays no biological, molecular parameters or transcriptional signatures have been established in the clinical practice.

Considering all these data published in literature however it has become evident that the investigation of molecular mechanisms related to resistance needs an integrated -omic approach, which have to consider genomic and transcriptional aspects of the disease, investigating not only coding and non-coding genes, but also known and unknown transcripts. These could lead to the identification of a complete molecular signature that is able predict response to Pt-compounds at time of patients' diagnosis and to understand the biological

features already present in primary tumour masses that will characterized the resistance disease, for whom no biological information are available at time of recurrence.

2. Aims

The high mortality rate that characterised HGS-EOC is mainly caused by the resistance that tumour sooner or later develops towards Pt-based compounds. Although hundreds of studies have been published over the last years in order to identify molecular biomarkers able to predict resistance to platinum-based drugs, none of these is currently used in clinical practice.

The identification of a biological signature associated to Pt response would help stratify patients' risk of relapse at time of diagnosis, thus avoiding unnecessary treatments to those patients who are intrinsically resistant towards Pt-based drugs.

To achieve this ambitious goal, the aim of this thesis is to contribute to the identification of a molecular signature associated with intrinsic platinum resistance, investigating the entire transcriptome and some genomic aspects in retrospective cohorts of cases opportunely selected as Pt-sensitive (Pt-s) and Pt-resistant (Pt-r).

As graphically summarised in figure 2.1, this study is subdivided into three main parts that are: pathway identification, genomic instability and transcriptome analysis. As it can be seen, the three steps are based on the use of different technological approaches. In fact, when this study begun the array technology was the gold standard to investigate gene expression profile. When NGS technology became popular and implemented in our laboratory, it allowed: *i*) an absolute quantification of the transcripts' abundance and a reconstruction of unknown or partially unknown transcripts and *ii*) to obtain data regarding Copy Number Alteration (CNA) and mutational profile. However due to the limits of access to NGS, RNA and DNA sequencing data were not obtained on the entire cohorts exploited for array-based experiments, but only in a subgroup of them.

Gene and miRNA expression data generated and confirmed using three independent cohorts of HGS-EOC patients (Materials and Methods section, paragraph 3.1), named as cohort A (n=99), B (n=143) and C (n=838, curatedOvarianData database) were integrated using an in house

developed algorithm (*micrographite*) to identify biological pathways driving Pt resistance and that could impact on patients' prognosis.

To integrate data obtained with mutational and more detailed transcriptional information, both DNA and RNA sequencing experiments were not performed on the entire cohorts A and B, but only on a sub-group of them named AI (n=28) and BI (n=55), best balanced in terms of Pt sensitivity (Results section, paragraph 4.3.1).

It is known from literature that the status of the genes belonging to the Homologous Recombination (HR) pathway can influence response to Pt compounds (98) so to test these evidences, using NGS approach DNA sequencing experiments were performed both in Cohort AI and BI (Figure 2.1). Data obtained were analysed in order to define, set and test an HRD score, able to discriminate HR proficient from HR deficient cases and to correlate this status with the response to Pt based chemotherapy.

In the last part of this study the entire transcriptome was sequenced (total RNA sequencing) to identify in cohort AI known or novel transcripts that are differentially expressed between Pt-sensitive and Pt-resistant HGS-EOC patients. RNA sequencing results obtained were then validated using a sequencing targeted approach both in cohort BI (technical validation) and in cohort AI (clinical validation).

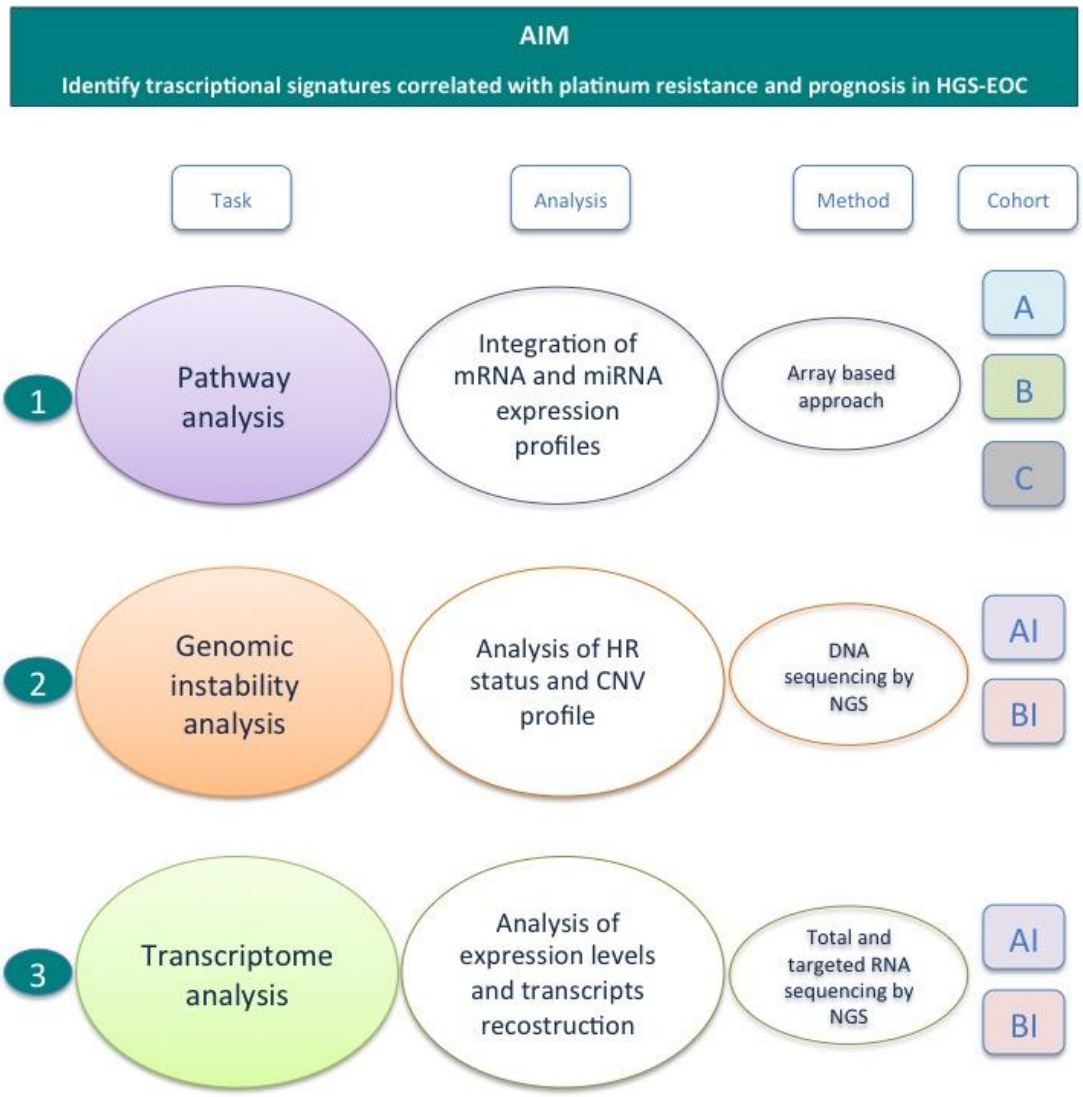


Figure 2.1 Study Overview: workflow of the study in which is briefly summarized the three main tasks faced in this thesis represented by the pathway analysis (1), the analysis of the genomic instability (2) and the analysis of the transcriptome (3). For each task the type of analysis performed and the technical method applied are reported. Moreover, for each of the three aspects investigated the cohorts (A, B, C, AI and BI) on which the experiments were performed are reported.

3. Materials and methods

3.1 Tissue Sample Collections

A total number of 242 HGS-EOC tumour samples, derived from two independent cohorts of patients (named A and B) were selected for this thesis work. All biopsies selected derived from HGS-EOC patients, naïve to chemotherapy, underwent debunking surgery and staging procedure according to the FIGO (International Federation of Gynecological and Obstetrics) guidelines. For this work only FIGO stage III/IV were selected. Based on the time lagging between the end of first line platinum-based therapy and relapse, patients from cohort A and B were classified into i) Pt-s, platinum-sensitive ii) Pt-ps, partially-sensitive and iii) Pt-r platinum-resistant as detailed described in the introduction (paragraph 1.3.3.2). Relapse after first-line platinum-based therapy was clinically identified (i.e. raise of CA-125 serum marker) and always radiologically confirmed with computed tomography (CT) scan. Data obtained from the analysis of cohorts A and B were tested *in silico* in another independent cohort (cohort C). In cohort C the time between the end of first line treatment and relapse is not available, thus, the stratification among Pt-r, Pt-s and Pt-ps was not possible.

Cohort A is composed by 99 snap-frozen HGS-EOC biopsies obtained at the Division of Obstetrics & Gynecology, ASST Spedali Civili, University of Brescia, between 2003 and 2013. Biopsies are stored in the biobank located at the "A. Nocivelli" Institute, ASST Spedali Civili of Brescia. A subset of this cohort (n=28, cohort AI) was selected balancing Pt-s and Pt-r cases, for Next Generation Sequencing experiments (see "Results" section paragraph 4.2.1).

Cohort B is composed by 143 snap-frozen tumor biopsies derived from HGS-EOC patients underwent debulking surgery at the Obstetrics and Gynecology Dept., San Gerardo Hospital (Monza, Italy). Biopsies are stored in the Pandora tumor tissue collection located at Mario Negri Institute for Pharmacological Research (Milan). A subset of this cohort (n=55, cohort BI) was selected on the basis to keep best balancing Pt-s and Pt-r cases, for Next Generation Sequencing

experiments (see “Results” section paragraph 4.2.1).

For both cohorts A and B clinical and anatomopathological information were registered and follow-up data were obtained from periodical gynecologic and oncological check-ups. Tumor samples were collected during cytoreductive surgery and frozen within 15 minutes in liquid nitrogen and stored long term at -80°C. The tumor content of the biopsies collected was evaluated with hematoxylin and eosin staining and only samples containing more than 70% of epithelial tumor cells were used for downstream analyses.

The local scientific ethical committees approved the collection and usage of all tumour samples and a written informed consensus was obtained from all the patients enrolled (study reference number NP1676 and NP1065). The study has been carried out following the Declaration of Helsinki (99).

Cohort C is composed by 838 samples obtained through the curated ovarian cancer dataset publicly available through the Bioconductor platform gathering together the TCGA data plus other ovarian datasets. The curated ovarian database as reported in the “curatedOvarianData” Bioconductor package was used as a first external independent validation set (Cohort C). The curatedOvarianData database (100) contains several normalized and batched-corrected ovarian datasets. Among these, we selected studies with platinum treated patients and complete follow-up for both progression free (PFS) and overall survival (OS). Unfortunately, the progression free interval (PFI) defined as the time between the relapse and the end of platinum-based therapy is not available. For this reason, samples of the curatedOvarianData cannot be stratified according to Pt-s and Pt-r. Here we will use PFS as a proxy of platinum resistance. The selection leads to 5 datasets (GSE30161, GSE9891, GSE49997, TCGA microarray and also TCGA.RNASeqV2) with a total of 838 samples.

3.2 Nucleic Acid extraction

3.2.1 RNA extraction

Total RNA enriched in miRNA fraction was extracted starting from about 30 mg of snap-frozen tumour tissue biopsies using miRNeasy mini kit (Qiagen®).

The RNA extraction procedure was subdivided in two main phases:

i) tissue homogeneization and cell lysis

ii) RNA purification

The homogeneization of snap frozen tissue was obtained through mechanic disruption (TissueLyser LT, Qiagen) using a phenol-guanidium thiocyanate solution (QIAzol Lysis Reagent, Qiagen®), which promotes the cell lysis and inhibits RNase activity. The subsequent addition of chloroform, followed by centrifugation (15 min at 12000g, 4°C), allowed to obtained three different phases: an upper colourless aqueous phase containing total RNA containing miRNA, a white interphase (DNA) and a lower red organic phase which contains lipids and proteins. The upper aqueous phase, after the addition of 1.5 volumes of 100% ethanol, which promotes RNA precipitation, was transferred into a silica membrane spin column (miRNeasy mini spin column, Qiagen®) for the purification steps that for this thesis work were performed automatically on QIAcube system (Qiagen®). The membrane's pores size allowed to retain all nucleotides > 18 base pairs (bp), so total not degraded RNA and miRNAs (22-25 bp). A series of washes with dedicated buffer followed by centrifugations purified the RNA extracted, and finally RNA enriched in miRNA fraction was eluted in 30 µl of RNase-free water.

The amount of RNA extracted was evaluated with a fluorimetric analysis (Qubit RNA BR assay kit, Thermo Fisher®) and the integrity quality control was performed through a commercial

automated capillary electrophoresis instrument (Tape Station 4200, Agilent Technologies®).

3.2.2 DNA extraction from snap-frozen tumour tissue biopsies

DNA was extracted from about 25-30 mg of snap-frozen tumour tissue biopsies using QIAamp DNA mini Kit (Qiagen®).

The protocol included the following steps:

- Chemical tissue lysis and enzymatic protein digestion of snap-frozen tumour tissue, exploiting dedicated SDS-buffer (ATL buffer) and Proteinase K, through an overnight incubation at 56°C.
- Addition of RNase A (100mg/ml) to remove RNA residual contamination
- Transfer of the material into a silica membrane spin column (QIAamp mini spin columns, Qiagen®) for the DNA purification steps that for this thesis work were performed automatically on QIAcube system (Qiagen®). In particular, DNA is adsorbed onto the QIAamp silica membrane during a brief centrifugation (800 rpm for 1 min), and following the washes with dedicated buffers ensured that proteins and other contaminants were not retained with DNA on the QIAamp membrane.
- Elution of purified DNA in 200 μl of distilled water.

The amount of DNA extracted was evaluated with a fluorimetric analysis (Qubit DNA assay kit, Thermo Fisher®) and the integrity quality control was performed through a commercial automated capillary electrophoresis instrument (Tape Station 4200, Agilent Technologies®).

3.2.3 DNA extraction from blood samples

Genomic DNA (gDNA) was extracted from whole-blood samples using Maxwell RSC Blood DNA Kit (Promega®). The protocol included the following steps:

- Cellular lysis and enzymatic protein digestion of 300 μl of whole-blood sample, through incubation (20 min at 56°C) of samples with lysis buffer and Proteinase K.
- Automatic purification of genomic DNA obtained using Maxwell RSC instrument (Promega®), which exploiting paramagnetic cellulose-based particles, allowed to capture and wash the gDNA extracted.
- Elution of purified gDNA in 50 μl of Elution Buffer.

Also in this case the amount of DNA was evaluated using fluorimetric assay (Qubit DNA assay kit, Thermo Fisher®) and the integrity of gDNA obtained was checked through a commercial automated capillary electrophoresis instrument (Tape Station 4200, Agilent Technologies®).

3.3 Pathway analysis

3.3.1 Gene and microRNA expression profile

Microarray mRNA and miRNA expression experiments were performed on the 99 HGS-EOC cases composing the cohort A, through microarray approach exploiting commercially available kits (G4851B human whole GE kit and G4470B Human miRNA kit, Agilent Technologies®), which allowed to investigate 26083 mRNA, 30606 lncRNA and 799 miRNAs.

Regarding mRNAs and lncRNA expression, the main steps are schematically reported in figure

3.1.

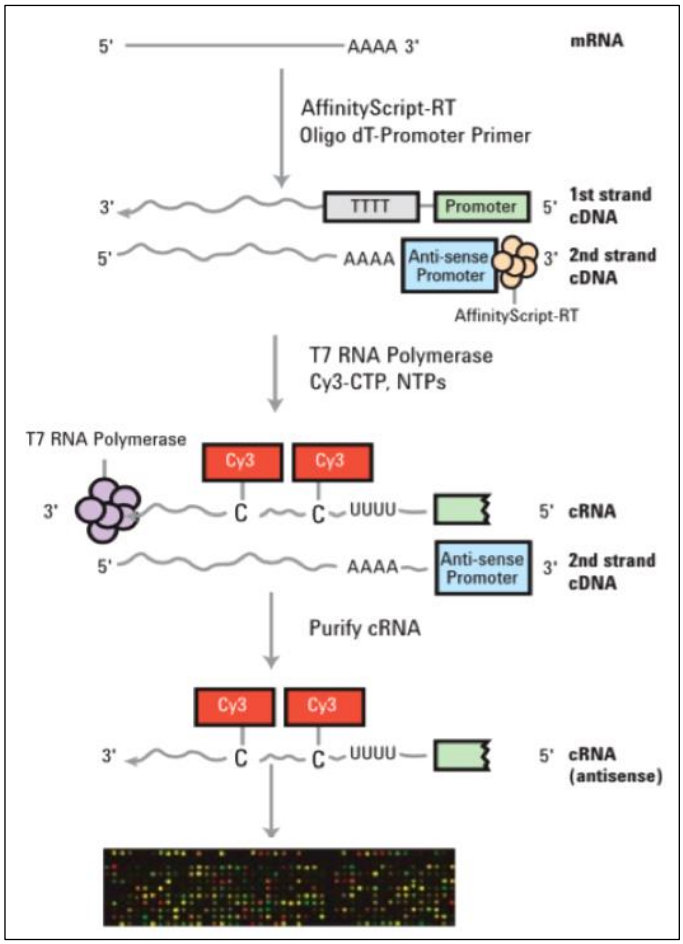


Figure 3.1 Schematic workflow of microarray experiment

Briefly, in gene and lncRNA expression microarray experiments 100 ng of total RNA were retrotranscribed in single strand cDNA exploiting reverse transcriptase enzyme in the presence of primer olig-dT bound to the poly-A tail of mRNA. Through a T7 RNA polymerase's activity, that recognized a specific sequence at 3' end of the second strand cDNA, the complementary RNA (cRNA) was synthetized and marked through the incorporation of cytosine conjugated with a fluorophore (Cyanine 3, Cy3). The conversion from cDNA to cRNA was necessary due to the major stability of DNA-RNA rather than DNA-DNA hybrids. In order to remove unbound Cy3 a purifiction step exploiting silica membrane spin columns (RNeasy mini kit, Qiagen®) was performed.

The concentration and the Cy3 incorporation efficiency in cRNA defined as "specific activity",

was calculated using NanoDrop ND-1000 (Thermo Fisher®) using the following formula:

(Cy3 concentration/cRNA concentration) x 1000 = pmol Cy3 for µg cRNA

At this point, cRNA was fragmented in order to prevent cross-hybridization or the formation of secondary structures which could affect the hybridization with probes. Once reaction was stopped through basic solution (pH 11), fragmented cRNA samples were put on the commercially available microarray glass slides, where were spotted clusters of 60 bp probes. The hybridization reaction between marked cRNA and spotted probes took place at 56°C for 17 hours with a rotation speed of 10 rpm.

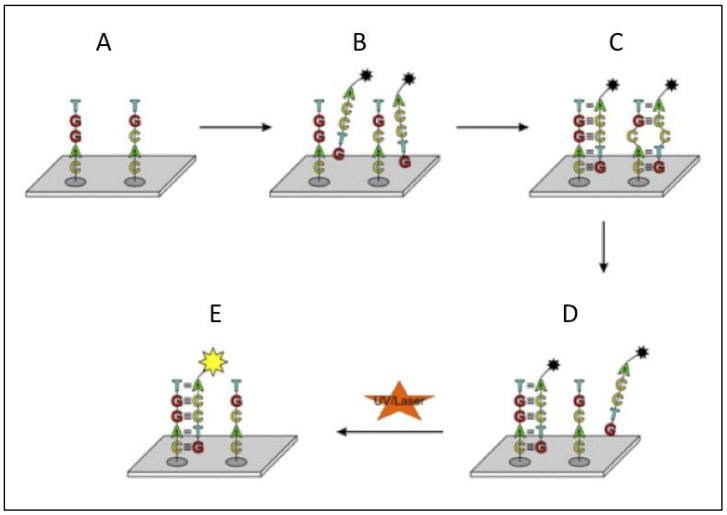


Figure 3.2 Schematic principle of oligonucleotide microarray. Two different probes were spotted on microarray slide (A). The hybridization of fluorophore (Cy3)-labeled cRNA with the immobilized probes (B) occured following the base paring Watson and Crick's model, resulting in a perfect and total pairing or in a mispair (C). After stringent washes which allowed to remove unbound or mispair labelled-cRNA sequences from the slide (D), fluorescence was detected after fluorophore excitation (λecc = 550 nm) (E). (modified from (101))

At the end of incubation slides were washed with dedicated solution.

Microarray slides were then inserted in G2565 Scanner (Agilent Technologies®) and exploiting the fluorescence properties of Cy3 (λecc = 550 nm and λem = 570 nm), the intensity of fluorescence signal due to the presence of hybridized marked (Cy3) cRNA was registered (figure 3.2).

A dedicated software (Feature Extraction version 11, Agilent Technologies®) was exploited to

associated fluorescence signal to the corresponding specific gene or lncRNA.

Regarding miRNA expression profile, as microRNA are characterized by a short nucleotide sequence (20-25 bp) it was necessary to develop a technical artefact to allow their specific hybridization on microarray slide. In particular, as schematically reported in figure 3.3, to 5' end of microarray probe (black sequence), was added by manufacturer a guanine (G) residue (black), complementary to Cy3-labeled pCp existing at 3' end of the target molecule (red), and an hairpin sequence (blue) aimed to stabilize the binding between probe-target and increase the size specificity.

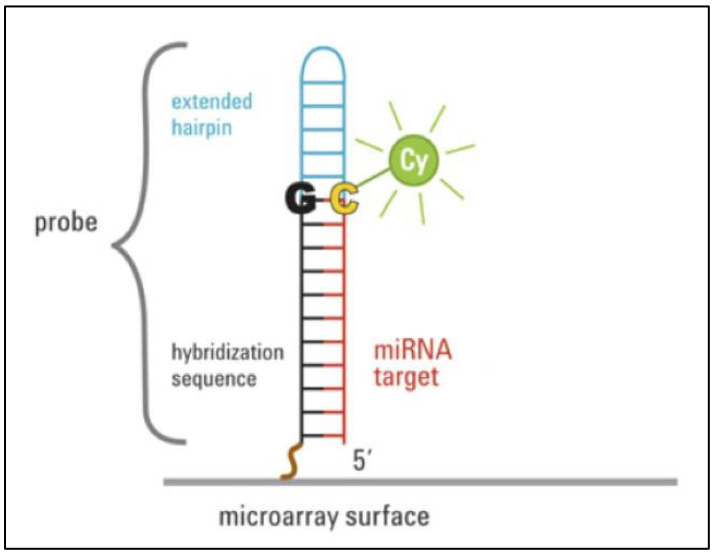


Figure 3.3 Probe design strategy to evaluate microRNA expression by microarray technology. In black is reported the probe's sequence that was characterized at 5' end by the presence of a guanine (G, black) and an hairpin sequence structure (blue) able to stabilized and to improve the specificity of probe-target (red) hybridization. Modified from (102).

Starting from 100 ng of total RNA containing microRNA fraction, samples were firstly dephosphorylated through Calf Intestinal Phospatase (CIP)'s activity (37°C, 30 min). After denaturation phase (100°C, 5-10 min), Cyanine 3 (Cy3) pCp was added to the 3' end exploiting T4 RNA ligase. After a purification phase on silica membrane spin columns to remove reaction reagents (i.e. DMSO) and unbound Cy3-pCps, samples were dried for 3 hours at 55°C using a vacuum concentrator and then resuspended in nuclease-free water. After a denaturation step, samples were put on the microarray slide and were hybridized, following the same principle

previously described, at 55°C for 20 hours with a rotation speed of 10 rpm. Slide washes and the fluorescence analysis performed occurred as described above for gene expression microarray experiments.

3.3.2 Microarray data processing

Data generated by microarray experiments, were analyzed by the bioinformatic team coordinated by Prof. Chiara Romualdi (Biology Department of University of Padova, Italy).

3.3.2.1 Filtering steps

Raw mRNA, lncRNA and miRNA expression signal was quantified using Agilent FeatureExtraction software. gProcessedSignal and gTotalGeneSignal were used as expression measures respectively for mRNAs and miRNAs. A first step of filtering was set up using the Agilent flag gIsPosAndSignif for mRNAs and lncRNAs and gIsGeneDetected for miRNAs. Probes with more than 60% of low-quality values were removed. A second step of filtering was performed on samples using Relative Log Expression (RLE) plot (103). Samples with RLE distribution markedly different from the general trend were removed from the following analyses. Probes mapping on the same gene annotation were averaged. Finally, data have been log transformed and quantile normalized. Coefficient of variation (CV) was used to select the most variable genes and miRNAs.

3.3.2.2 RUV normalization

In data with large scale gene expression studies, the observations are commonly contaminated by sources of unwanted variation such as platforms or batch effects. These unwanted variations during analysis can lead to weak associations and to missing important signals. Using RUV normalization (104) negative control genes and replicate samples was used to estimate unwanted variation data. The proposed method was translated on microarray data and implemented in the bioconductor package RUVnormalize (105). They generally manage to remove unwanted variation without losing the signal of interest and compare favorably to state-of-the-art corrections. RUVnormalize, using positive and negative control probes (n=48), was applied (with number of normalizing factor k=2) to remove unwanted variation and batch effect. Hereafter, differentially expressed (DEs) coding and non-coding elements between Pt-r and Pt-s were identified using permutational moderated t-test, as implemented in samr R package (permutational number set to 1000).

3.3.2.3 Network analysis Micrographite

In order to integrate miRNAs in pathway analysis, micrographite (106) (107) gave us a lot of advantages: (i) the possibility to integrate and (ii) analyze miRNA and mRNA expression profiles using pathway information and (iii) to biologically contextualize miRNA–mRNA validated and predicted interactions. In our study, micrographite was used to identify integrated circuits of mRNAs and miRNAs associated to the therapy response (Pt-r vs Pt-s).

Pathway topologies derived from *graphite* (108), a Bioconductor package developed to store, manage and convert pathway annotations into gene–gene networks. *graphite* is a pathway data interpreter that, following biologically driven rules, is able to solve the complexity of the pathway modules to generate interaction networks suitable for topological pathway analyses. KEGG pathways as available through *graphite* (108) and miRNA-target gene interactions

available on “TargetScanHuman” database release 7.1 (109) were considered for the analysis. Only miRNA-target interactions characterized by a Pearson correlation coefficient $|r| \geq 0.4$ and *q*-value ≤ 0.05 were included into pathway annotation. Then, the topological pathway analyses used in micrographite are a modified version of CliPPER (106) (110). CliPPER is a Bioconductor package that implements a topological pathway. This approach is based on two steps. In the first step, pathway graphs were compared in terms of means and variance between groups. On these selected pathways, it identifies portions of the pathway mostly associated with the phenotype. Into details, path identification is based on the graph decomposition into small-connected components, called cliques. Each clique is tested independently (according to the test on the means and/or concentration matrices) and then a significant level (p-value) for each clique is obtained, creating a path as a list of adjacent significant cliques. Clipper is able to identify and score all of these paths in the graph. Paths obtained have a score that is a function of all the p-values of the cliques contributing to the path and to the higher score correspond the better path. Then, the upper-scored 10th percentile of the portion of significant pathways (in mean and variance with $p<0.1$ and 10 000 permutation) are combined into a non-redundant meta-pathway. Finally, the portion of the meta-pathway mostly associated to the phenotype is revealed.

3.3.2.4 Network analysis: meta-pathway extension

A significant portion of genes is not annotated into KEGG pathways, and then is not considered by micrographite pipeline. In absence of this information, we considered the construction of networks having all the protein-protein interactions using computational methods for signalling pathways and protein complex identification in specific diseases. Studies have also shown that proteins with larger number of interactions can include families of enzymes, transcription factors, and intrinsically disordered proteins, among others.

Thus, to extend our network taking into consideration the excluded genes, we used STRING

(111) and BioGRID (112) databases: differentially expressed genes without KEGG annotation were added to the network if at least one of their interactors (as reported by STRING and BioGRID) were present in the network. Cytoscape (113) was used to visualize the integrated regulatory network.

3.3.3 Gene retrotrascription and qRT-PCR validation

The validation of gene expression data obtained from microarray experiments was performed in 242 HGS-EOC cases (cohort A and cohort B) exploiting an independent technique, the quantitative real-time PCR (qRT-PCR).

Firstly, up to 10 μg of total RNA obtained from samples (see paragraph 3.2.1) was retrotranscribed in cDNA, using commercially available kit (High Capacity cDNA Archive Kit, Applied Biosystem®). The retrotranscription phase consisted in the addiction to each RNA sample of a master mix composed by a reverse transcriptase enzyme, deoxyriboNucleotide TriPhosphates (dNTPs), random primers able to bound the polyA tail of mRNAs and a dedicated buffer, followed by a thermacycler incubation (25°C for 10 min, 37°C for 120 min). RNA retrotranscription was followed by qRT-PCR reaction that was prepared in 384 well reaction plates followed the manufacturer's indications (QuantiFast SYBR Green, Qiagen®) exploiting an automatic liquid handling system (EpiMotion 5075LH, Eppendorf®). qRT-PCR reaction was prepared with reagents indicated in table 3.1.

Reagents	Final concentration
2x QuantiFast SYBR Green PCR Master Mix	1X
Primer FW (100 μM)	1 μM
Primer RV (100 μM)	1μM
cDNA	< 100 ng
Nuclease-free Water	To adjust final volume

Table 3.1 Components of qRT-PCR reaction. QuantiFast SYBR Green Master Mix included the HotStar DNA polymerase, SYBR Green dye molecules, dNTPs and reaction buffer containing KCl and $(NH_4)_2SO_4$ to optimize reaction's condition. FW, forward. RV, reverse.

Briefly, The qRT-PCR reaction allows the amplification of specific cDNA sequences through the repetition (40 cycles) of three fundamental steps: i) denaturation of ds cDNA by heat (95°C for 10 sec) ii) the annealing of gene-specific forward and reverse primers able to bound complementary sequences flanking DNA target region, which permit the addition of oligonucleotides into the rising DNA strand and iii) the extension of the new DNA strand through a DNA polymerase activity (annealing and extension phases were performed simultaneously at 60°C for 30 sec) (figure 3.4).

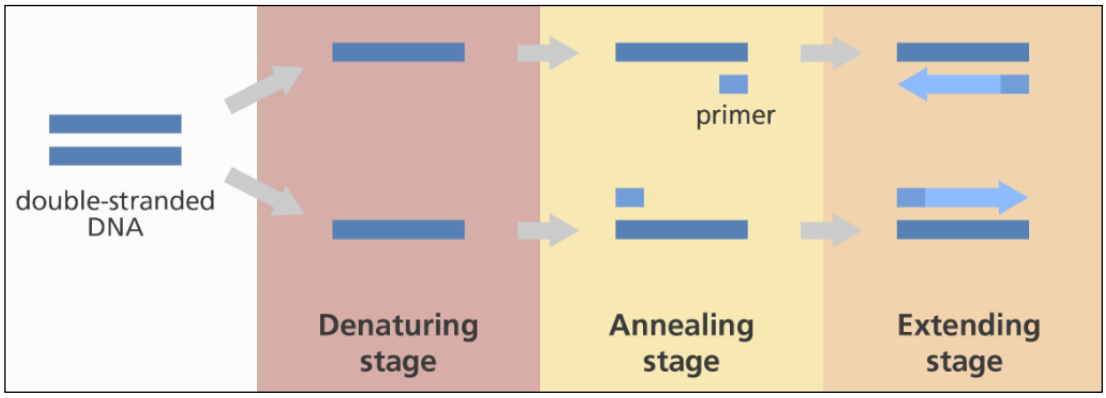


Figure 3.4 qRT-PCR cycle. In figure are reported the three fundamental steps composing each qRT-PCR cycle: the denaturation, the annealing and the extension phase.

During the polymerization phase SYBR Green dye molecules (λecc = 488 nm and λem = 522 nm) became intercalated within the double strand DNA product. When excited SYBR Green

intercalated within dsDNA was able to emit a fluorescence signal 1000 over then the one emitted from the free unbound SYBR Green, representing so an optimal indicator of total DNA product obtained during each PCR cycle.

Results obtained at the end of a qRT-PCR experiments were represented by an amplification curve (figure 3.5) where was manually set a threshold that intersected the amplification curve at a specific point, threshold cycle (C_T), representing the level of a significant increase in fluorescence due to DNA amplification. For each gene, C_T values obtained were normalized by subtracting to the mean value of specific sample the mean value of housekeeping genes (ΔC_T).

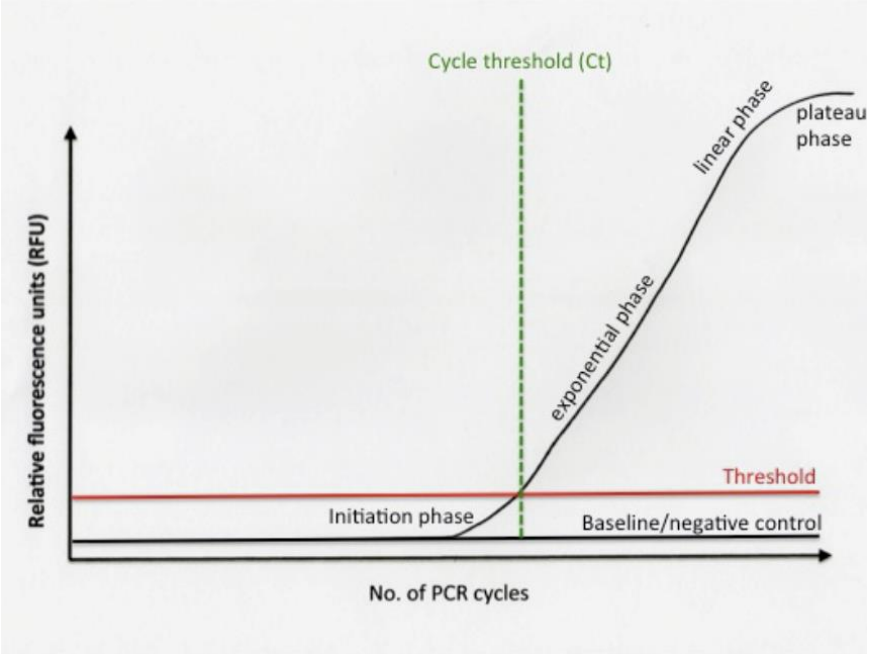


Figure 3.5 Amplification curve of a qRT-PCR experiment. On x-axis is reported the number of PCR cycles, on y-axis the relative fluorescence units (RFU) registered at the end of each PCR cycle. Threshold manually set allowed to distinguish the fluorescence due to background signal from fluorescence related to DNA amplification.

qRT-PCR experiments reported in this thesis work were performed using Applied Biosystem® 7900 instrument. Each one of the 242 samples analysed was run in triplicate and three housekeeping genes (GAPDH, HPRT1 and PPIA) were used for the normalization of data obtained.

The list of primers used, which were design using Primer-blast (https://www.ncbi.nlm.nih.gov/tools/primer-blast/) and Primer3 (http://primer3.ut.ee/)

platforms are reported in table 3.2.

Gene	FW primer sequence	RV primer sequence	TA (°C)	Amplicon Lenght
GAPDH	AGGTCGGAGTCAACGGATTT	TTAAAAGCAGCCCTGGTGAC	60	58
HPRT1	TGAATACTTCAGGGATTTGAATCAT	CTCATCTTAGGCTTTGTATTTTGC	60	76
PPIA	GCGTCTCCTTTGAGCTGTTT	CCTTTCTCTCCAGTGCTCAGA	60	79
NOTCH2	GAAGGCAGGTCTCCTGTGTC	ATCTTCTGTGCAGTCAGCCC	60	145
HLA-A	ATGAAGGCCCACTCACAGAC	GTGAGAACCGTCCTCGCTC	60	81
TOPORS	GAAGAAATAGGGCCTTTCCG	TGCCATTATCATGAAGCCAGT	60	128
EYA4	GTAAACCAGTTTGAAAAATGTTCTGT	AATAGCCGAAAACCCACTTT	60	50
NPY	GGGGATTTTCCCTTG	AAAACCAAAATGTCTTTCTCTCCA	60	50
FANCI	TACGGGTAACGGAAGTGTGG	TCACAGAACTCCGCCACAAA	60	70
FANCA	GAAGAGGCCTTCCTGCATGT	GGTTGCCCTGACCCTTGAG	60	127
PPP1CA	GACCGTGGCGTCTCTTTTAC	TCTTCTACCACCTGGTGTGCT	60	101
PIGR	AGAGGCAGGGGTTACCAACT	TCCTGTGCAATGTTTTAGCCAC	60	89
HLA-F	GCTGCAGTGTGAGACAGCTT	TGTATGTTCGTGAGGCACAA	60	87
CACNA1C	CGTGGCTGCTCCTCCTATTA	CATAGTTGGAACCTTGGTGGTT	60	97
CREB3	CTTTCTGAGGTACCGAGCGA	GAGAATGTTCAACGACGCTG	60	81
STAT1	AACCTCGACAGTCTTGGCAC	GAGACATCCTGCCACCTTGT	60	96
BAD1	TGTGGACTCCTTTAAGAAGGGAC	CACCAGGACTGGAAGACTCG	60	102
FLCN	GGACCGGATCTACCTCATCA	AACGCCAAACTGGGAGAAG	60	103
APOL6	TTTCTCCAGCCCAGACACTC	TCAAATGATTTTCTTCTCTCCACG	60	139
ROCK2	GGTGATCGTATTCTTCCAGTGA	TTTTGGGCCATCATATTTCA	60	51
ETV7	CAAGCCAGATGTGAAGCTCA	CTGGATGCGGAGTCTTCCT	60	78
PRKG1	GCAGATCCAGGAGATTGTGG	CAACCTTACCATCTTCCATGACA	60	116
SDF2L1	GTCCAACAACCAGGAGGTGA	TGACGGAACACAGAGGTGC	60	145
GCH1	GGTTGAAGCAACACACATGTG	CCAACGCACACACACTGAAT	60	164
PPP1R12A	AAGGTGAAGTTCGACGATGG	TCATCAATGCAAGCCTGGT	60	149
CREB5	CTAAGTCCAGAGAGTAGCCCTCC	CCAGGTCTGATGGTGCATT	60	187

Table 3.2 Sequences of qRT-PCR reverse and forward primers. For each gene validated are reported the sequences of forward (FW) and reverse (RV) primer, the annealing temperature of primers (TA) and the amplicon length generated.

Data obtained by qRT-PCR experiments (ΔC_T) were analysed using "R" platform by bioinformatics and statistic group coordinated by Prof. Chiara Romualdi (University of Padova, Italy). In particular, T-test was used to test the mean difference between Pt-r and Pt-s samples on log transformed qRT-PCR expression values.

3.3.4 Statistical analysis

Statistical analysis reported in this thesis work were entirely performed by statistic group located at Dept. Biology of the University of Padova (Italy).

Kaplan-Meyer curves were used to visualize patients' survival. Univariate analysis was performed with the log rank test, multivariate analysis was performed using Cox proportional hazard model as implemented in *survival* R package. Overall survival (OS) was calculated considering the time lagging between the diagnosis and the death for any cause or the last follow-up. Instead, progression free survival (PFS) is defined as the time from diagnosis to disease progression or last follow-up. The optimal cutoffs of gene expression values were estimated using i) the maximally selected rank statistics as implemented in the *survMisc* R package for survival analysis and ii) the maximum value of specificity and sensitivity using ROC curves for prediction of Pt-r and PT-s patients. Residual tumour and age were used as covariates in the multivariate analysis. Samples with missing survival data were excluded from the analysis. Results are reported as p-value, hazard ratio (HR) and 95% confidence intervals (CI). Multivariate logistic regression has been performed using the *glm* and *brglm* R packages while odd ratio statistics have been obtained using *fmsb* R package.

P-values combination in cohort C has been performed using the Tippett's method as implemented in *metap* R package.

The combination of *SIII* genes has been performed separately for cohort A and B using the

standardized expression profiles within cohort to minimize batch effect. All the analyses were performed using R platform.

3.4 Genomic instability analysis

3.4.1 Next generation sequencing of tumour and genomic DNA

Next Generation sequencing experiments on DNA extracted from snap-frozen tumour tissue samples (tumour DNA, tDNA) and where available from matched whole-blood samples (genomic DNA, gDNA) were performed exploiting OneSeq Constitutional Research panel (Agilent Technologies®), which allow studying genome-wild Copy Number Variations (CNVs) and LOH, and evaluating the presence of indel and single nucleotide variants (SNVs) in 6000 disease-associated genes defined by ClinGen (https://clinicalgenome.org) (figure 3.6).

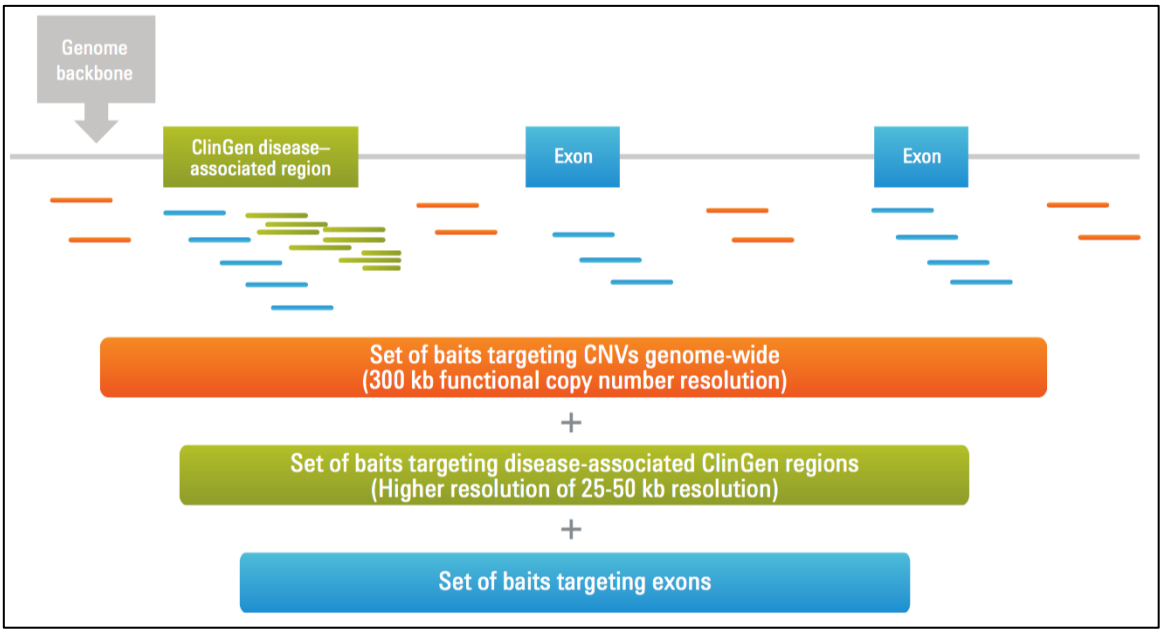


Figure 3.6 Schema design of OneSeq Constitutional Research Panel (Agilent Technologies®). OneSeq targets a functional copy number resolution of 300 kilo bases (kb) in the genome-wide backbone, with an even higher resolution of 25-50 kb targeted in disease-associated regions defined by ClinGen. Moreover, the panel targets exonic regions of 6000 disease-related genes for detection of indels or point mutations at high reads depths.

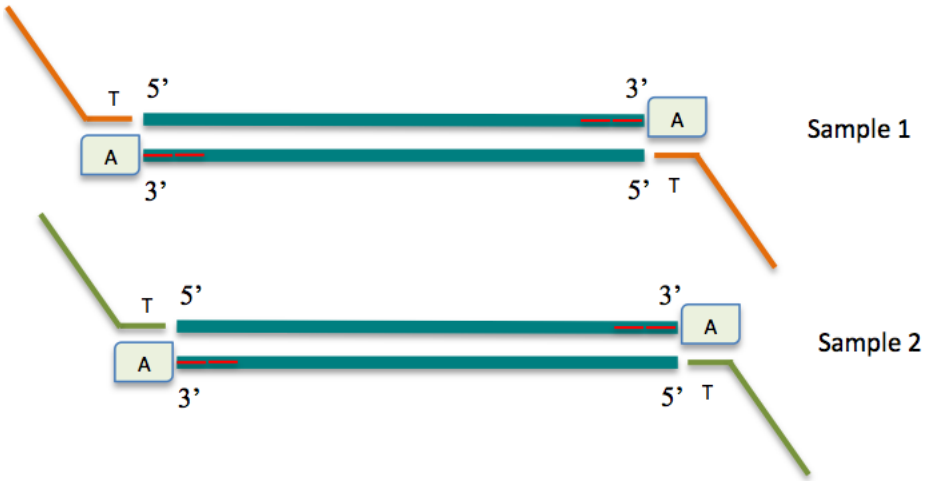
Following the manufacturer's instruction, the following steps were performed for libraries construction:

1- Mechanic shearing of tumour and genomic DNA, that was performed exploiting Bioruptor ultrasonicator (Diagenode®). In particular 200 ng of purified DNA in a total volume of 100 µl diluted in TRIS10mM pH8 was sheared by ultrasonication (30 sec ON + 90 sec OFF for seven cycles) and purified using 1.8X AMPure XP beads (Beckman Coulter®). After elution in distilled water, the fragment size of sheared DNA was checked using Tape Station 4200 (Agilent Technologies®). DNA fragment size peak was expected between 200-300 bp.

2- End repair, which consisted in the blunt of both 3' and 5' end trough the activity of a DNA polymerase. At the end of this step a purification of sample with AMPure XP beads (Beckman Coulter®) was required.

3- Adenylation of 3’ end of the DNA fragments, in which dATP at 3’ was added at 3’ ends of the blunt DNA fragments to prevent concatemerization and to allow the ligation of adapters with complementary dT overhangs. Also at the end of this step, purification with AMPure XP beads (Beckman Coulter®) to remove reaction components was performed.

4- Adapter ligation, which through DNA ligase activity, allowed assigning to each sample a specific and unique barcode sequence (8 bp). Exploiting AMPure XP beads (Beckman Coulter®), purification of samples was performed at the end of this step.



5- Pre-capture PCR amplification in order to amplify the adaptor ligated libraries prepared, following the thermal cycler program reported in the protocol. Purification of libraries at the end of PCR reaction was required.

6- Assessment of quantity (Qubit®) and quality (Tape Station 4200, Agilent Technologies®) of libraries obtained. The fragment size peaks were between 225 to 275 bp, as expected.

7- Pooling of 15 library samples with different barcode, to obtain a 750 ng DNA in 3.4 µl of volume.

8- Overnight hybridization of pool (denaturation step followed by 65°C for 16-20 hours, lid 105°C) with commercial biotinylated-probes (see the beginning of this paragraph)

9- Capture of the biotinylated-probes that bound during the night the specific sequences targets, through streptavidin beads (figure 2.7).

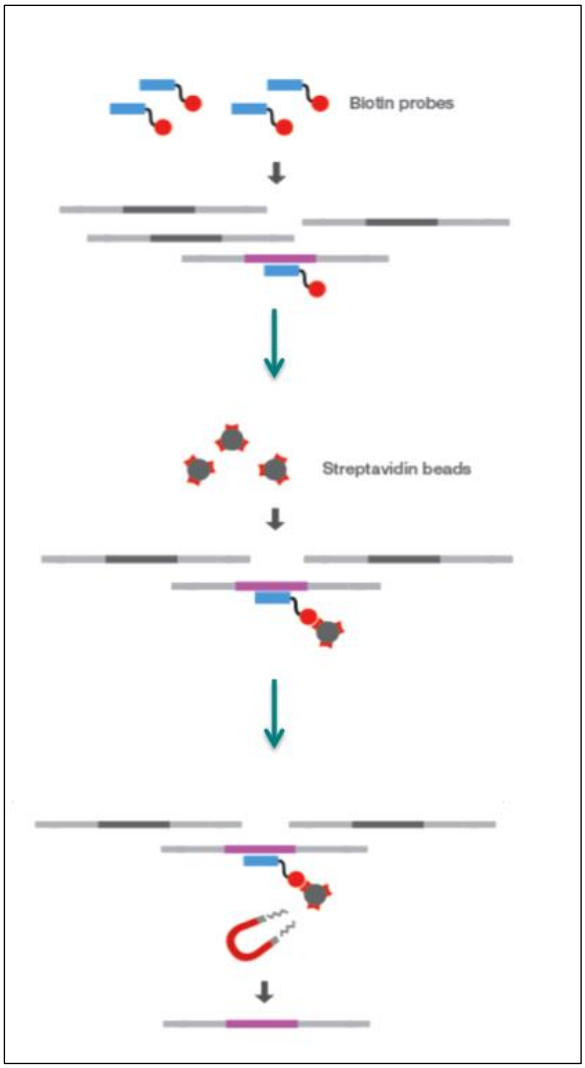


Figure 3.7 Capture-based enrichment. Bound of biotinylated probes with target cDNA sequence, and subsequent interaction between probes-cDNA with streptavidin beads.

10- Amplification of captured libraries followed by purification of them using AMPure XP beads (Beckman Coulter®).

11- Qubit® quantification and evaluation of the quality of libraries obtained performed exploiting Tape Station 4200 (Agilent Technologies®). As expected, the average fragment length obtained was between 250-350 bp, with a peak at approximately 300 bp.

Considering the capacity of a single flow cell and the average coverage necessary for the specific aims (i.e. analysis of germline or somatic mutations) 15 sample libraries per run were sequenced on NextSeq-500 sequencing system (Illumina®) using High-output 300-cycles or Mid-output 300-cycles cartridge/flow cell in the cases of libraries derived from tumour DNA or genomic DNA respectively.

3.4.2 DNA sequencing data analysis

Data generated by DNA sequencing experiments, were analyzed by the bioinformatic team coordinated by dr. Sergio Marchini at Dept. of Oncology of the Institute.

3.4.2.1 Quality control and alignment to the genome

Raw sequenced fragments from the instruments (reads) were first of all subject to quality control. Samples were checked for the average length of the reads versus the expected lengths (150 bp) and for the quality of the base calls (Q30, percentage of bases with a quality score equal or above 30, that is 99.9% confidence of a correct base call).

After the initial quality control, raw reads were aligned to the human reference genome, build 19 (hg19), using the Burrows-Wheeler Aligner (BWA) (114). At this point bases with low

quality were masked and excluded from subsequent analysis, without actually removing them from the data ("soft clipping").

The PCR steps used in library preparations can cause a percentage of artifacts (ranging from 5 to 10% of the total data) show up in the data as duplicate, identical reads: thus, during the alignment, duplicate reads from artifacts were removed.

3.4.2.2 Germline and somatic variant calling and copy number alteration estimation

The aligned data from blood samples were used to initially estimate germline variants occurring in the *BRCA1* and *BRCA2* genes. To this aim, the Genome Analysis Toolkit (GATK;(115)), version 4, was used to call gemline variants, using the Haplotype Caller.

For each variant, the chromosomal position, the reference base, the alternate (variant) base, the sequencing depth, the genotype and the fraction of variant versus reference reads were reported. The resulting variants were then annotated to add information on the gene they were falling in, the effect on the protein, and prevalence in the general population: this was done with the Variant Effect Predictor (VEP; (116), a tool which queries the Ensembl database and associated data sources (such as Uniprot and ClinVar) to obtain detailed annotation.

Germline variants were screened in the *BRCA1* and *BRCA2* genes with these criteria: minimum allelic fraction (ratio of variant reads versus normal reads) of at least 25%; presence in the general population less than 1% (as reported in the 1000 genomes data set and the gnomAD population database); unambiguous definition of pathogenicity from ClinVar (that is, only variants classified as "pathogenic" or "likely pathogenic").

Subsequently, the blood samples were used to build a reference data set used for both somatic variant calling in tumor biopsies (to remove normal variants occurring in the population) and for somatic copy number variant detection, to provide a robust baseline. Samples with detected *BRCA1/BRCA2* germline variants were not included in the reference.

MuTect 2 (117) was used to generate raw variant calls from the blood samples, which were then processed to only include variants occurring at least in five or more samples, following the best practices for variant calling outlined by the Broad Institute. This call set was then used as reference for the subsequent analyses.

With regards to copy number, a pooled reference was generated with CNVkit (118). Briefly, the reads from the blood samples were divided into "bins" of roughly 200bp in size (average length of an exon) and counted for each bin. Then, information from public databases and from the genome sequence was used to correct the numbers for GC content and for mappability (as the sequence content of some genomic regions may be inherently harder to align to opposed to others).

Once the reference data were ready, we performed the variant calling and copy number analysis on the tumor samples. For the former case, we used two different variant calling software:

- MuTect 2
- VarDict (119)

As reference, we used both the matched blood sample (if available) and the pooled reference we built. Variants were then annotated with the VEP as described above, and collected into a database (GEMINI; (120)) for further annotation and analysis.

Copy number analysis was performed with CNVkit, using the pooled reference previously built. Reads were binned and counted in the same manner as the blood samples, then bin counts were compared to the reference to generate copy number ratios, which were then converted to log base 2, after removing all the bins which were too noisy or with too low coverage (< 4X).

As the bins were by nature rather small, we collapsed them into segments of equal copy number with a process called "circular binary segmentation" (CBS;(121)).

3.4.2.3 Purity and ploidy estimation

In order to estimate *in silico* the tumor content and the ploidy of each sample, we used an approach which combined both the state of heterozygous SNPs across the whole genome and the genome-wide copy number. These in turn were fed to a mathematical and probabilistic model to infer purity, ploidy, and the state of loss of heterozygosity (LOH).

To this aim we used PureCN (122) which used the calls from VarDict, including the calls marked as germline, and the copy number data from CNVkit. Firstly, we built a reference in the same fashion as the previous analyses, using the copy number and the germline variant data from the pooled blood samples. Subsequently, each sample was analyzed with PureCN: the software used an iterative approach of searching the most likely ploidy / purity combination from the copy number and variant data, followed by optimizing and fitting until the most likely solution was found.

The analysis yielded estimated purity and ploidy, LOH state across the whole genome, purity and ploidy corrected copy number calls and variants. Each sample was also checked for noisy data or other problematic metrics and flagged specifically if so.

3.4.2.4 Variant interpretation

Variants from Mutect2, VarDict, and PureCN were filtered to remove low depth of coverage calls (likely sequencing artifacts) and then filtered to remove common polymorphisms, synonymous variants, and likely variants of unknown significance using a "prioritization" approach which measures the presence of a specific variant in common sequence databases such as Ensembl, gnomAD, and the 1000 genomes project. Variants with fractions lower than 5% were also removed.

The resulting calls were screened for variants in genes of interest and candidate variants were then manually confirmed by viewing the wild type versus variant reads for each locus of interest using a genome-wide variant viewer, the Integrated Genome Browser (IGV; (123).

3.4.2.5 Homologous Recombination Deficiency (HRD) score calculation

We used the copy number data and the variant data from PureCN to calculate the homologous recombination deficiency (HRD) score as described by Telli et al. (124) and as implemented by Sztupinszki et al. (125). To do so, we extracted from the PureCN data the copy number values for each allele for every variant identified: the copy number of the major allele ("major copy number") and the copy number for the minor allele ("minor copy number").

We then built segments of homogeneous total (major + minor) copy number from the individual points. At this point, we calculated the three measures which are part of HRD:

- Loss of heterozygosity (LOH): count of every segment with LOH (minor copy number = 0) if it is greater or equal than 15 Mbp;

- Telomeric allelic imbalance (TAI): Count of contiguous segments which start from the telomeric region up to the centromere (but not including the whole chromosome) that exhibit allelic imbalance between the major and minor alleles;

- Large scale transitions (LST): The number of copy number transitions (that is, when two adjacent segments have different copy number) between segments of at least 10Mbp in length, not spanning the centromere.

The final HRD score was calculated as the simple algebraic sum of LOH, TAI and LST. The thresholds for HRD were made in a group of 30 high quality samples (predicted tumor fraction greater than 30%, not flagged for bad data quality) comparing the distribution of the HRD score between the germline *BRCA1/2* mutated samples (the true positives) with the controls. The threshold was set as the 5^{th} percentile of the distribution of the HRD scores in *BRCA1/2* mutated

samples. Samples were called HR deficient if their score was greater than the threshold, otherwise they were called HR proficient.

3.5 RNA sequencing

3.5.1 Total RNA sequencing

Total RNA sequencing experiments were performed using TruSeq stranded total RNA kit (Illumina®).

The first step for RNA sequencing, was the synthesis of a cDNA library from total RNA. This step is mandatory because RNA molecules can not be directly sequenced due to their low chemical stability.

The major steps in library preparation, schematically represents in figure 3.8, involves the following:

1) Preparation of a total RNA input ranging from 0.1 to 1 µg, considering the Qubit quantification.
2) Deplation of rRNA from total RNA (Ribo-Zero Gold kit, Illumina). This step was done by annealing the total RNA to oligo-dT magnetic beads. As rRNA is the most abundant component of RNA (80-90%) its depletion was necessary in order to focus on more informative parts of the transcriptome. At the end of this step a clean-up was performed.
3) Enzymatic fragmentation of the purify RNA, in order to break the RNA strands into multiple small fragments

4) Reverse-transcription of the RNA fragments using a reverse transcriptase, and random hexamer primers, thus producing the first strand cDNA. The addition of Actinomycin D to the reaction mix, inhibits DNA-dependent synthesis, while allowing RNA-dependent synthesis, improving strand specificity.
5) Syntheses of the second/opposite strand of the cDNA and remove the RNA template. This strand was generated incorporating dUTP in place of dTTP. The incorporation of dUTP quanches the second strand during amplification, because the polymerase did not incorporate past this nucleotide. After the synthesis a clean-up was performed in order to separate the ds cDNA from the second strand reaction reagents. At the end of this process a blunt-ended double strand cDNA was obtained.
6) Adenylation of 3' ends of the blunt fragments to prevent them from the ligating to one another during the adapter ligation reaction. A corresponding single "T" nucleotide on the 3' end of the adapter provides a complementary overhang for ligating the adapter to the fragment
7) Ligation of adapters to both ends of the ds cDNA. Each adapter had a six-nucleotide difference in adaptor sequence. The use of a different index for each library reaction allowed to pool libraries later for sequencing, yet still allowing for tracing the sequence back to the original library based on the adaptor sequence. At the end of this step a clean-up was performed.
8) Enrichment of the library by polymerase chain reaction (PCR) amplification using sequences from the adaptor as primers. At the end of this step clean-up was performed to remove PCR reagents.

9) Evaluation of library obtained. In order to Qubit (Termo Fisher®) quantification of DNA library templates the quality control analysis, which evaluate by a capillary electrophoresis the size distribution of the libraries (the final product should be a band at approximately 260 bp)

10) Normalization procedures on the basis of the amounts of DNA and pooling of libraries generated.

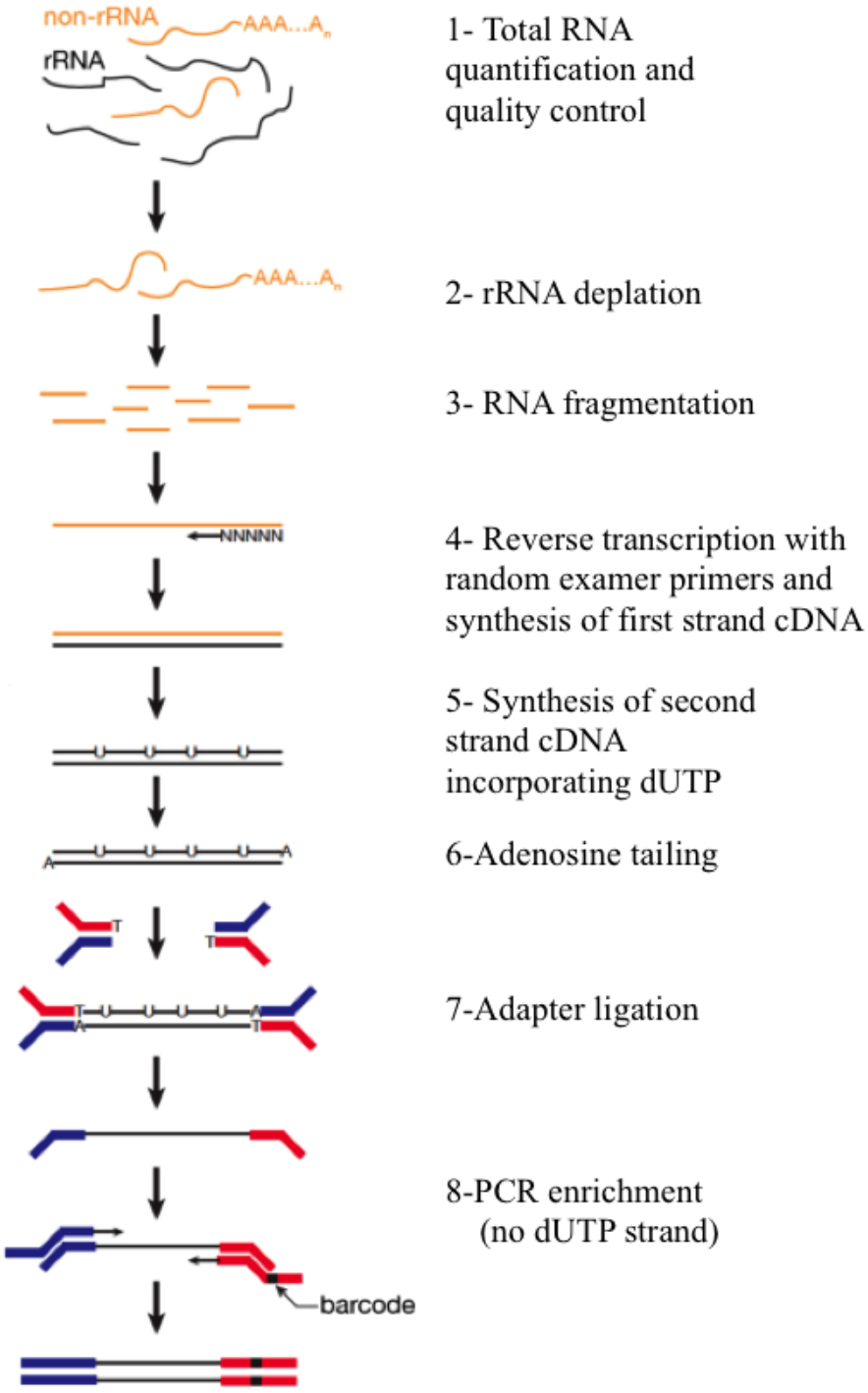


Figure 3.8 Workflow of RNA sequencing library preparation. Modified from (126).

In order to obtain at least 80 millions of reads for patient, eight sample libraries per run were sequenced outsourcing (Personal Genomic S.r.l. spin-off of the University of Verona, Verona Italy) on Hiseq 2500 Ultra High-Troughput Sequencing System (Illumina®).

3.5.2 Total RNA sequencing data analysis

Data generated by RNA sequencing runs, were analyzed by the bioinformatic team coordinated by Prof. Chiara Romualdi (Biology Department of University of Padova, Italy).

The major steps of the differential expression analysis performed are schematically reported in figure 3.9.

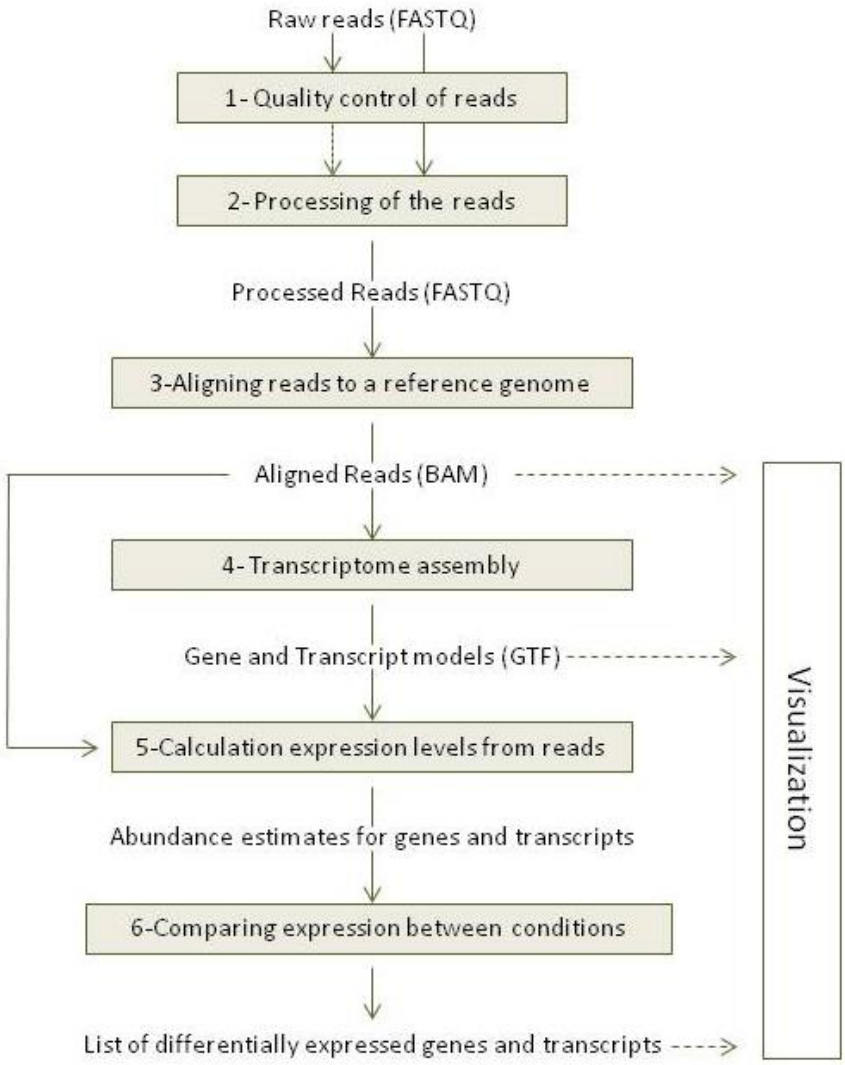


Figure 3.9 Workflow of the differential expression analysis for annotated transcripts

On the basis of figure 3.9 a brief description of each step is reported above.

Step 1, quality control of the reads: in the first step a general quality control analysis was performed automatically on the Illumina® platform. This look at the overall quality of the reads

(in FASTQ format, that is an extension of the standard FASTA format for sequences with the addition of a confidence score, called Phred score, for each called base). Reads were scanned i.e. for low-confidence bases, biased nucleotide composition, adapters, duplicates. The output of this step was basic statistics such as the number of reads and the quality information, which guides the pre-processing, decisions in the subsequent step.

Step 2, pre-processing of the reads: the goal of this step was to remove low-quality bases and artefacts such as adapter sequences from individual reads. Experimental artefacts were also removed. For example, poly A tails were removed since they interfere with analysis steps later. Reads were also trimmed to remove adaptors or low quality bases. After pre-processing, the data were in a cleaned and polished form and they were able to undergo the next data analysis steps. Also this step was automatically performed on the Illumina® platform.

Step 3, alignment of reads to a reference genome: the goal of this step was to find the point of origin for each read. To this aim, in this project a two-pass alignment strategy using STAR (Spliced Transcripts Alignment to a Reference) package was exploited (127). Reads were mapped to a reference genome sequence (*H. Sapiens*, hg38.p2 version) and annotated (Ensembl 79). The output of this step was an alignment file (Binary Alignment File; BAM) for each patients, in which were indicated the mapped reads and their related positions in the reference.

Step 4, transcriptome assembly: in this project this step was performed using StringTie (PMID: 25690850) (128), that is able reconstruct and identify automatically for each patient known and unknown transcripts (.gtf form file), giving for each one information as the start and the end position, confidence score or feature type (i.e. exon, transcript, mRNA). Cuffmerge and cuffcompare (PMID: 20436464) were respectively used to create a consensus transcriptome (output: a single merged.gtf file starting from the .gtf files previously generated) and to compare the reconstructed transcripts with existing annotations.

Step 5, calculating transcripts' expression levels: in order to estimate the transcripts abundance, in this project RSEM (RSEM version 1.2.21) was exploited (129). The .gtf file previously obtained (step 4) was used to map all the reads generated (FASTQ files) on reconstructed

transcripts. For each transcript and for each patient the number of raw reads count was calculated. Transcripts with less than 10 counts in at least 60% of samples per class were filtered out.

Step 6, comparing gene expression between conditions: before the comparative analysis, a normalization procedure is required because of possible differences in read numbers and transcriptome composition. In this project data were normalized using full quantile normalization as implemented in the edgeR (130) software package (R version 3.2, Bioconductor version 3.2, edgeR version 3.12.0). This tool calculated for class a mean value of each transcript identified and considered as differentially expressed those with a corrected p-value (q-value or false discovery rate, FDR) < 0.05.

3.5.3 Pathway analysis of known genes identified as differentially expressed in total RNA sequencing

Known genes were tested with functional enrichment with REACTOME (https://reactome.org) that is an open-source, open access, manually curated and peer-reviewed pathway database.
The p-values resulted from the test were corrected for multiple testing using the False Discovery Rate (FDR) as transcripts could not be considered as independent entities. The pathways reconstructed by DAVID were called significant, whether its corrected p-value (q-value or FDR) did not exceed 0.05.

3.5.4 Targeted RNA sequencing

A custom targeted RNA sequencing panel (SeqCap RNA enrichment System, Roche), designed on the basis of the differentially expressed transcripts identified, was used to confirm total RNA sequencing results.

The steps for the construction of targeted RNA sequencing libraries are quite similar to those that characterized the preparation of the total RNA sequencing ones (paragraph 3.5.1). In fact, starting from 100 ng of purified total RNA (input: 10-1000 ng), RNA underwent to enzymatic fragmentation in order to break the RNA strands in smaller fragments, and directly the synthesis of the 1st and 2nd strands cDNA was performed. After a clean-up double stranded cDNA aimed to eliminate the 2nd strand reaction reagents, the adenylation at 3'end of blunt fragments to prevent concatemerization and to allow the ligation of the adapters with complementary dT overhangs occurred. After the ligation of adapters, that were formed by six nucleotides and difference one from each other for nucleotide sequence, a double clean-up was performed. The bound of sample-specific adapters for each library reaction allowed pooling 7-10 for the subsequent hybridization capture. In fact, after a pre-capture PCR amplification and a clean-up aimed to eliminate PCR reagents, a overnight hybridization (95°C for 10 min + 47°C for 16-20 hours) with custom biotinylated-probes designed to cover and capture all the 1730 transcripts identified as differentially expressed in paragraph 3.5.2 was performed. Once ended the hybridization phase, the biotinylated-probes that bound during the night the specific sequences targets, were isolated using streptavidin beads. In fact, streptavidin beads were able to capture and isolated the biotinylated probes plus bound target cDNA.

After subsequent washes with dedicated buffers to eliminate reaction reagents and unbound fragments, streptavidin beads-biotinylated probes with bound target cDNA underwent PCR reaction as reported in manufacturer's user's guide. The PCR allowed i) to release the target cDNA from the beads-probles complex (the denaturation step: 98°C for 45 sec), and ii) amplified the target cDNA captured. At the end, after a final clean-up and the elution of purified

target fragments, Qubit quantification evaluation of the quality of libraries obtained was performed exploiting Tape Station 4200 (Agilent Technologies®). As expected, the average fragment length obtained was between 150-500 bp, with a peak at approximately 300 bp.

Considering the capacity of a single flow cell (high-output 300 cycles, Illumina®) and the barcoding procedures applied during the libraries preparation, seven (for cohort AI, see Results section, paragraph 4.2.1) and ten (for cohort BI, Results section, see paragraph 4.2.1) libraries per run were sequenced on Nextseq Sequencing System (Illumina®).

3.5.6 Targeted RNA sequencing analysis in cohort AI

Also in this case, data generated by targeted RNA sequencing were analysed by bioinformatics group coordinated by Prof. Chiara Romualdi (Biology Department of University of Padova, Italy).

After quality control and processing of reads generated, that occurred automatically on sequencing system, in order to assess the target capture strategy efficacy, reads (FASTQ file) were aligned on the genome sequence (hg38.p2 version) and annotated (Ensemble 79) using STAR package (as described in paragraph 3.5.2, "Step 3"). Transcripts quantification was then obtained aligning the reads on the reconstructed transcript sequences derived from the deep total RNA sequencing analysis. This process was performed using Salmon (PMID: 28263959) that is a tool for wicked-fast transcripts quantification, which requires a set of target transcripts either from a reference or from *de novo* assembly (FASTAQ files), as in this case, to quantify (131). In this RNA sequencing analysis transcripts with less than ten counts in at least 60% of samples per class were filtered out. In order to reduce artefacts and biases that affect the resulting expression measures, data obtained were normalized exploiting EDA (Exploratory Data Analysis) Seq (PMID: 22177264) (104). The EDAseq tool is based on a within-lane normalization followed by a between-lane normalization procedure. The within-lane (or intra-sample) normalization procedure adjusts samples' data for their GC-content effect, basing on

the principle that CG-content differs between genes but not between samples. The between-lane normalization instead, corrects for distributional differences in libraries/samples that can affect the counts and lead to variation (i.e. sequencing depth).

Using the normalized data, differential expression analysis was computed using edgeR (R version 3.2, Bioconductor version 3.2, edge R version 3.12.0) and only transcripts with a corrected p-value (q-value o FDR) <0.05 were considered as differentially expressed between pt-r and pt-s classes.

A *in silico* depths dilution, simulating different sequencing depths starting from 40 millions reads for sample, was performed through multiple (n=1000) read resamplings approach in order to set up the sequency depth optimization for the targeted RNA sequencing in cohort BI which represented the clinical validation of results obtained. This test allowed to estabilish a threshold of 20 million of reads per sample to obtain robust and riproducible RNA sequencing results.

3.5.7 Targeted RNA sequencing analysis in cohort BI

Targeted RNA sequencing in cohort BI was performed multiplexing 10 samples per run to have an average 20 millions of paired reads per sample.

Transcript quantifications have been performed with Salmon (PMID: 28263959), using the reconstructed transcript sequences derived from the deep total RNAseq analyses. Transcripts with less than 5 counts in at least 30% of samples per class were filtered out.

Expression data normalization was carried out with the package EDASeq (PMID: 22177264) (see paragraph 3.5.6). Differential expression was computed using edgeR (R version 3.2, Bioconductor version 3.2, edge R version 3.12.0) and transcripts with a corrected p-value (q-value or FDR) <0.05 were considered differentially expressed.

4. Results

4.1 Pathway analysis

In the first part of this work, three independent cohorts of patients were analysed to identify a regulatory network, composed from both coding and non-coding genes, shaping the different biology of Pt-s and Pt-r cases and to investigate its association to prognosis in terms of Overall Survival (OS) and Progression Free Survival (PFS).

In figure 4.1 is reported a schematic workflow of the study. In the first part of this work, using data obtained from array experiments in cohort A, an integrated approach considering both coding and non-coding genes were exploited to identify a regulatory network associated with Pt response. In order to test the robustness of data obtained 23 out of 136 elements were selected and tested using an independent technique (qRT-PCR) in cohort A and B. Confirmed the association of these elements with Pt response, their prognostic value (overall survival, OS and progression free survival, PFS) were evaluated and tested both in cohort A+B and in a larger cohort of HGS-EOC patients that is considered the benchmark for HGS-EOC (cohort C).

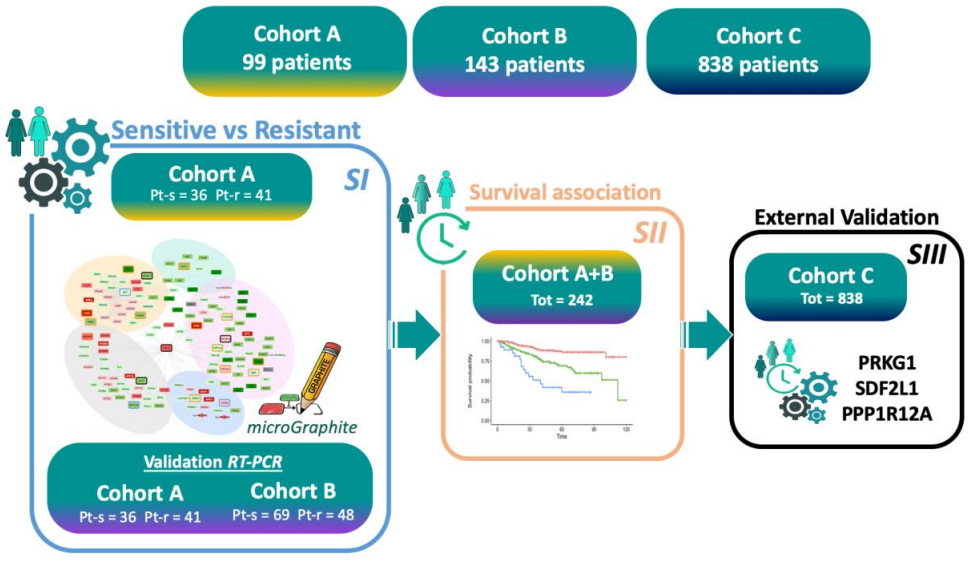


Figure 4.1 Workflow of the pathway analysis of the study

4.1.1 Cohorts' description

To the aim of this part of the work, a total number of 242 ovarian cancer biopsies derived from two independent cohorts (cohorts A and B) of patients, were selected for the aim of this study. Results obtained from the analysis of cohort A and B were later tested *in silico* in a larger cohort of patients (cohort C). Clinical-pathological features of HGS-EOC patients belonging to cohort A (n=99), B (n=143) and C (n=838, Curated Ovarian Cancer dataset), are shown in table 4.1. All the biopsies selected were collected at primary surgery, before chemotherapy, from patients with diagnosis of HGS-EOC stage III and IV, according to FIGO (Federation International of Gynecology and Ostetrics) criteria. The median age at diagnosis was 63 years (cohort A), 57 years (cohort B) and 59 years (cohort C), respectively. The median follow-up time was 3 years for cohorts A and B, and 2 years for cohort C. Demographic and disease characteristics reported for cohort A are comparable to those reported in literature for HGS-EOC patients (cohort C), whereas for cohort B they are characterized by longer OS and PFS compared to the HGS-EOC benchmark (cohort C).

Differently from that reported in the cohort C, cohort A and B are enriched in patients (60% both in cohort A and B and 34% in cohort C, respectively) with suboptimal residual tumour (RT), that is defined as a tumour mass > 1 cm, that was not possible to eradicate from the abdominal cavity with the surgery. This discrepancy could be due to two main reasons: i) RT is subjectively determined by surgeons during the surgical procedure and ii) the cohort C cases were enrolled from 90's, when an optimal debulking surgery was considered when RT<2 cm. Regarding the Pt response, both in cohort A and B patients were best balanced on the basis of their Pt sensitivity. In particular, patients defined as Pt-s, Pt-r, and partially sensitive (Pt-ps) were 36%, 42% and 19% in cohort A, respectively. In cohort B they were 48%, 34% and 17%, respectively. Unfortunately, information regarding the Pt Free Interval (PFI) is not available for cohort C, thus patients have been indirectly classified into Pt-s and Pt-r using Progression Free Survival (PFS) measure, in particular patients were defined Pt-s with PFS>12 months and Pt-r

with PFS<6 months. Being PFS as a rough estimation of PFI, data regarding the number of Pt-s and Pt-r patients, is not reported in table 4.1.

In the table 4.1 are also reported the clinical-pathological information relative to cohorts AI and BI, that are sub-cohort of the larger cohorts A and B. Cohort AI and BI, used for both DNA and RNA sequencing experiments are detailed described in paragraph 4.2.1.

Clinical Annotations	Patients' characteristics					
	Cohort A	Cohort B	Cohort C	Cohort AI	Cohort BI	
	No. Of Patients, N(%)	No. Of Patients, N(%)	No. Of Patients, N(%)	No. Of Patients, N(%)	No. Of Patients, N(%)	
Total No. Of patients	99	143	838	28	55	
Median (range) years	63 (36-85)	55 (33-83)	59 (22-89)	66 (36-85)	54 (33-79)	
Histotype						
Serous	99 (100%)	143 (100%)	838 (100%)	14 (100%)	55 (100%)	
FIGO Classification: stage, substage						
III, A	2 (2%)	0 (0%)	12 (1,5%)	1 (4%)	0 (0%)	
III, B	1 (1%)	5 (3%)	45 (5%)	0 (0%)	0 (0%)	
III, C	74 (75%)	118 (83%)	670 (79%)	19 (67%)	47 (85%)	
III, undetermined	-	-	5 (0,5%)	1 (4%)	0 (0%)	
IV	22 (22%)	20 (14%)	106 (14%)	7 (25%)	8 (15%)	
Platinum Status[$]						
Sensitive (PFS≥12 months)	36 (36%)	69 (48%)	-	14 (50%)	30 (55%)	
Partially Sensitive (6<PFS<12)	19 (19%)	25 (17%)	-	0 (0%)	0 (0%)	
Resistant (PFS≤6 months)	41 (42%)	48 (34%)	-	14 (50%)	25 (45%)	
NA	3 (3%)	1 (1%)	-	-	-	
Median follow up (range) years	3 (0-13)	3 (0-19)	2 (0-17)	2 (0-11)	4 (0-10)	
Median PFS (range) months[*]	14,4 (1,2-142)	31 (0,8-239)	14,3 (0,3-166)	10,9 (14,6-63,3)	20,4 (3,6-128,7)	
Median OS (range) months	36,2 (1,15-161)	53 (2-239)	28,1 (0,3-214)	31,8 (7,8-137,3)	48,5 (4,7-128,7)	
Residual Tumor						
R0 (RT=0)	20 (20%)	26 (18%)	-	2 (7%)	8 (15%)	
R1 (0<RT≤1)	20 (20%)	31 (22%)	-	8 (29%)	9 (16%)	
R2 (RT>1)	59 (60%)	86 (60%)	-	18 (64%)	38 (69%)	
Optimal	-	-	557 (66%)	-	-	
Suboptimal	-	-	281 (34%)	-	-	

Table 4.1 Clinical and demographic description of cases enrolled in cohorts A, B and C. PFS, Progression Free Survival. OS, Overall Survival. RT, Residual Tumour. In cohort C the Pt status is not available. * Calculated on 95 patients (4 patients do not have PFS follow-up).

4.1.2 Identification of a regulatory network associated with response to therapy

Overall data reported in literature have shown that a single gene approach signatures failed to stratifying HGS-EOC patients' risk of relapse at time of diagnosis, a compendium of transcriptomic data, including both coding and non-coding genes, was generated to characterize the different biology of Pt-s and Pt-r tumours. Pt-s (n=36) and Pt-r (n=41) cases have been retrospectively selected from cohort A and, on the basis of array experiments, mRNA, miRNA and lncRNA profiles analysed using *micrographite* algorithm, that allows to move from a list of single genes towards a comprehensive map of functionally related networks. *Micrographite* analysis resulted in a network composed of 131 mRNAs and 5 miRNAs (Figure 4.2, Table 4.2). The network wired five different major functional processes that are: i) transcription regulation, ii) transmembrane ion transport (such as calcium ion transport and sodium ion export), iii) cell cycle regulation and response to damage, iv) fatty acid metabolism, v) antigen presentation. The above five functional processes are interconnected by PRKG1 gene, which acts as hub of the network allowing crosstalk among the five different functional areas (Figure 4.2).

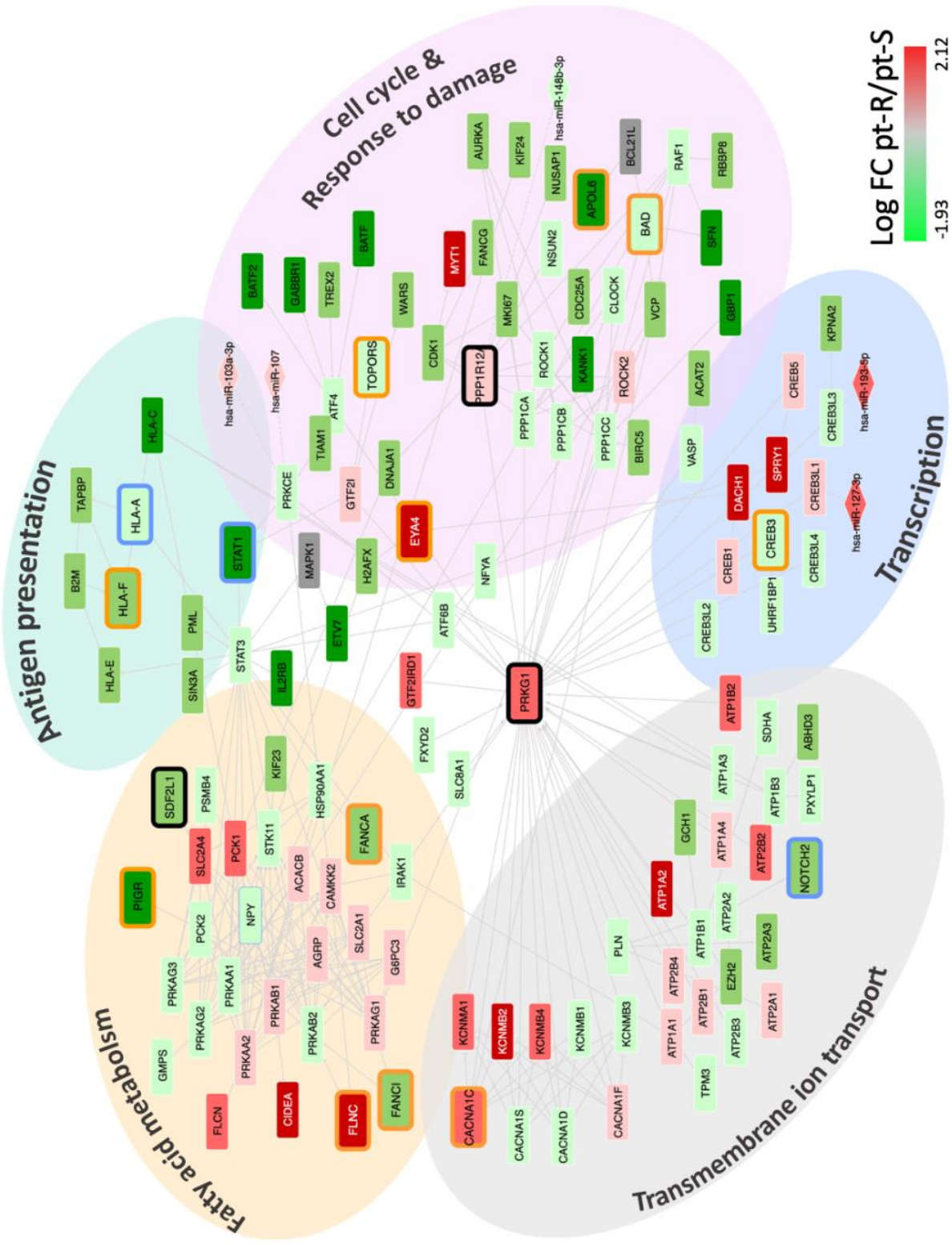


Figure 4.2 Regulatory network associated to Pt response. Integrated (mRNA and miRNA) network obtained through the comparison of Pt-s and Pt-r samples in cohort A, using pathway-based approach. The colours of the nodes represent the log Fold Change (log FC) Pt-r/Pt-s.

As the cell network interactions can be characterized not only by great changes in expression levels of genes (differentially expressed genes, DEG) but also by fine regulations well connected one to each other, *micrographite* algorithm selected not only DEG or genes with significant logFC value (≤ 0.5 or ≥ 0.5) but also genes and miRNAs with interconnection functions or with *in silico* predicted interactions.

Tables 4.2 A, B, C, D, E and F reports the 131 genes and five miRNAs composing the network, subdivided on the basis of their role in one or more functional processes identified. Three functional areas are particularly enriched in elements of the network that are the i) fatty acid metabolism (29 elements, 21%) ii) cell cycle and response to damage (46 elements, 34%) and iii) transmembrane ion transport (33 elements, 24%) process. The antigen presentation and transcription processes (Table 4.2 B and D) are identified by the presence of respectively 10 (7%) and 14 (10%) elements. Moreover, there are another eight genes that do not belong to any specific functional area, but represents nodes and the hub (PRKG1) of the network (Table 4.2 F).

A

Functional Area	Gene Name	Validation criteria		
		DEG	logFC	Literature evidence
Fatty acid metabolism	FANCI	0.029		
	FLNC	0.029	1.3	-
	CIDEA			
	FLCN			
	GMPS			
	PRKAG2			
	PRKAA2			
	PRKAB1			
	PRKAB2			
	PRKAG1			
	PRKAG3			
	PRKAA1			
	AGRP			
	SLC2A1			
	G6PC3			
	IRAK1			
	FANCA	0.017		
	CAMKK2			

	ACACB			
	NPY	-	-	Medeiros PG et al., Int J Cancer 2012 Medeiros et al., Peptides, 2013
	PCK2			
	PIGR	0.038	-1.93	
	SLC2A4			
	PCK1			
	STK11			
	HSP90AA1			
	KIF23			
	PSMB4			
	SDF2L1	0.064		

B

Functional Area	Gene Name	Validation criteria		
		DEG	logFC	Literature evidence
Antigen presentation	SIN3A			
	PML			
	HLA-E			
	HLA-F	0.017	-	-
	HLA-A	0.11	-	-
	B2M			
	TAPBP			
	HLAC			
	STAT1	<0.001	-	-
	has-mir-103a-3p			

C

Functional Area	Gene Name	Validation criteria		
		DEG	logFC	Literature evidence
	H2AFX			
	MAPK1			
	PRKCE			
	TIAM1			
	GTF2I			
	DNAJA1			
	EYA4	0.077	2.1	-
	NFYA			
	ATF4			

Cell cycle and Response to damage	has-mir-103a-3p			
	has-mir-107			
	BATF2			
	GABBR1			
	TREX2			
	TOPORS	0.11		
	BATF			
	WARS			
	CDK1			
	PPP1R12A	-	-	Serrano I. et al., Nature Comm 2013 Zhang C. et al., Mediators of inflamm 2015
	PPP1CA	-	-	Wang et al., Plos One 2011 Xia et al., Plos One 2014
	PPP1CB			
	PPP1CC			
	BIRC5			
	ACAT2			
	VASP			
	ROCK2	-	-	Zhong et al., Oncogene 2018
	ROCK1			
	KANK1			
	MKI67			
	MYT1			
	FANCG			
	AURCA			
	KIF24			
	NSUN2			
	NUSAP1			
	CDC25A			
	CLOCK			
	VCP			
	GBP1			
	SFN			
	BAD	-	-	Douglas C. et al., Clin Canc Res 2011
	APOL6	<0.001	-	-
	has-mir-148B-3P			
	BCL21L			
	RAF1			
	RBBP8			

D

Functional Area	Gene Name	Validation criteria		
		DEG	logFC	Literature evidence
Transcription	ACAT2			
	CREB3L2			
	CREB1			

	CREB3	-	-	Shanyang He, et al., Oncol Lett 2017
	DACH1			
	SPRY1			
	UHRF1BP1			
	CREB3L4			
	CREB3L1			
	CREB5	-	-	ProteinAtlas
	CREB3L3			
	KPNA2			
	has-mir-127-3p			
	has-mir-193-5p			

E

Functional Area	Gene Name	Validation criteria		
		DEG	logFC	Literature evidence
Transmembrane ion transport	CACNA1C		0.5	-
	CACNA1S			
	CACNA1D			
	CACNA1F			
	KCNMA1			
	KCNMB2			
	KCNMB4			
	KCNMB1			
	KCNMB3			
	PLN			
	ATP1A1			
	ATP2B4			
	TPM3			
	ATP2B1			
	ATP1B1			
	ATP2B3			
	EZH2			
	ATP2A2			
	ATP2A1			
	ATP2A3			
	ATP1A2			
	GCH1	0.053	-0.5	-
	ATP1A4			
	GCH1			
	ATP1A4			
	ATP1A3			
	ATP1B2			

	ATP2B2			
	ATP1B3			
	SDHA			
	NOTCH2	0.077	-0.5	-
	PXYLP1			
	ABHD3			

F

Functional Area	Gene Name	Validation criteria		
		DEG	logFC	Literature evidence
Hub	PRKG1		0.8	Hub of the network
Interconnection nodes	STAT3			
	IL2RB			
	ETV7	<0.001	-1.08	
	SLC8A1			
	FXYD2			
	GTF2IRD1			
	ATF6B			

Table 4.2 A, B, C, D, E and F. Elements of the network. In tables 4.2 are reported the name of genes and miRNA composing the network subdivided for functional area: fatty acid metabolism (table 4.2A), antigen presentation (table 4.2B), cell cycle and response to damage (table 4.2C), transcription (table 4.2D), transmembrane and ion transport (table 4.2E). Table 4.2F reported the genes representing nodes or hub of the network. Genes highlighted in yellow colour are those for which orthogonal validation was performed (see paragraph 4.1.3). The criteria for the selection are reported: DEG (q-value), logFC (Fold Change) and literature evidence (First author, Journal, Year of publication).

4.1.3 Network validation

To demonstrate the robustness and the reproducibility of the identified network, 23 out of 136 elements of the circuit were arbitrary selected and their different expression orthogonally validated by an independent technique (qRT-PCR), in both cohort A (technical validation) and in cohort B (biological validation). The 23 elements have been selected on the bases of at least one of those three criteria:

i) DEG (q-value<0.15)

ii) log fold change greater than ± 0.5

iii) literature evidence

The genes selected for orthogonal validation are highlighted in yellow in tables 4.2 A, B, C, D, E, and F and for each one the principle for the selection is reported. Data obtained in cohort A confirmed 20 out of 23 elements (87%) of the network as differentially expressed between Pt-s and Pt-r patients (Table 4.3). GCH1, NPY and PPP1CA have not been confirmed as differentially expressed.

	COHORT A										
	Gene	t-test pvalue	mean Pt-r	mean Pt-s	Odd ratio	Lower CI(95%) Odd ratio	Upper CI(95%) Odd ratio	Odd ratio pvalue	Pred. Pt-r (%)	Pred. Pt-s (%)	Threshold
Up-regulated in Pt-r	CACNA1C	**<0.001**	3.61	1.54	**1153.40**	53.57	24834.69	**<0.001**	100.00	95.12	6.17
	ETV7	**<0.001**	4.92	2.62	**25.58**	7.68	85.19	**<0.001**	86.11	82.50	21.41
	EYA4	**<0.001**	4.47	3.35	**18.83**	6.07	58.44	**<0.001**	77.78	85.71	13.82
	FLNC	**<0.001**	3.46	1.55	**53.75**	13.50	213.96	**<0.001**	91.67	85.71	6.63
	GCH1	0.56	2.64	2.40	**2.32**	0.93	5.78	0.07	54.29	66.67	4.21
	NPY	0.58	-0.03	-0.38	**2.06**	0.79	5.36	0.14	47.06	70.27	0.61
	PPPR12A	**<0.001**	4.49	2.21	**497.73**	26.56	9327.85	**<0.001**	100.00	88.10	10.69
	PRKG1	**<0.001**	3.82	1.95	**267.10**	37.41	1907.04	**<0.001**	97.22	92.86	9.91
	STAT1	**<0.001**	2.62	1.07	**105.29**	22.30	497.18	**<0.001**	91.67	92.68	3.55
Up-regulated in Pt-s	APOL6	**<0.001**	1.86	3.46	**801.29**	39.99	16056.57	**<0.001**	100.00	92.86	6.90
	BAD1	**<0.001**	1.76	3.86	**40.15**	11.02	146.19	**<0.001**	83.33	90.48	9.62
	CREB3	**<0.001**	1.98	3.14	**31.99**	9.33	109.67	**<0.001**	83.33	88.10	6.65
	CREB5	**<0.001**	0.51	3.43	**56.01**	12.66	247.79	**<0.001**	94.44	80.95	3.01
	FANCA	**<0.001**	1.09	2.91	**61.79**	15.43	247.47	**<0.001**	88.89	90.48	4.67
	FANCI	**<0.001**	1.91	3.36	**52.96**	13.29	211.06	**<0.001**	83.33	92.86	6.90
	HLA-A	**<0.001**	-0.55	3.13	**223.56**	36.55	1367.28	**<0.001**	94.44	95.24	3.01
	HLA-F	**<0.001**	1.53	2.79	**26.35**	8.03	86.51	**<0.001**	83.33	85.71	4.60
	NOTCH2	**<0.001**	2.01	4.16	**43.95**	11.34	170.40	**<0.001**	91.43	83.33	9.66
	PIGR	**<0.001**	1.11	2.53	**14.08**	4.70	42.19	**<0.001**	83.33	75.61	3.63
	PPP1CA	0.94	3.60	3.61	**2.17**	0.81	5.81	0.12	77.78	39.02	9.67
	ROCK2	**<0.001**	2.08	3.38	**8.96**	3.25	24.70	**<0.001**	75.00	76.19	7.16
	SDF2L1	**<0.001**	0.62	2.99	**48.66**	11.17	212.08	**<0.001**	94.44	78.57	3.29
	TOPORS	**<0.001**	1.14	3.00	**115.00**	22.90	577.46	**<0.001**	94.44	90.24	4.66

Table 4.3 Differential expression evaluated by qRT-PCR between Pt-r and Pt-s samples of the 23 selected genes in cohort A. For each gene, the mean average expression value of Pt-s and Pt-r samples and t-test p-value is reported. In cohort A the thresholds are the best expression cut-off that maximize specificity and sensitivity and Pred Pt-s/Pt-r (%) is the percentage of predicted Pt-s and Pt-r patients obtained using these thresholds. In gray are reported those genes that were not validated. Table is divided in up and down regulated genes in order to have odd ratios always greater than 1: for genes up-regulated in Pt-r patients the reference group is Pt-r, for genes up-regulated in Pt-s the reference group is Pt-s.

Biological validation experiments in cohort B confirmed differences in the expression value between Pt-s and Pt-r cases for 17 out of the 20 genes (85%) genes (Table 4.4).

	COHORT B						
	Gene	**t-test**	**mean Pt-r**	**mean Pt-s**	**Pred.**	**Pred.**	**Accuracy (%)**
Up-regulated in Pt-r	**CACNA1C**	**<0.001**	3.23	1.00	53.06	92.65	76.07
	ETV7	0.56	1.09	1.25	2.04	97.06	57.26
	EYA4	**<0.001**	5.52	3.87	87.76	39.71	59.83
	FLNC	**<0.001**	4.90	2.23	79.59	75.00	76.92
	GCH1	0.87	1.68	1.73	30.61	63.24	49.57
	NPY	0.88	1.10	1.04	100.00	0.00	41.88
	PPPR12A	**<0.001**	6.14	3.16	83.67	66.18	73.50
	PRKG1	**<0.001**	4.80	2.18	73.47	75.00	74.36
	STAT1	**0.02**	5.96	5.51	100.00	0.00	41.88
Up-regulated in Pt-s	**APOL6**	**<0.001**	3.30	5.27	42.86	98.53	75.21
	BAD1	**<0.001**	4.81	5.92	8.16	98.53	60.68
	CREB3	**<0.001**	3.40	4.81	20.41	98.53	65.81
	CREB5	**0.99**	6.80	4.64	4.08	88.24	52.99
	FANCA	**<0.001**	2.25	3.58	48.98	75.00	64.10
	FANCI	**<0.001**	3.63	5.32	28.57	94.12	66.67
	HLA-A	**<0.001**	5.19	6.60	0.00	100.00	58.12
	HLA-F	**<0.001**	3.48	5.18	16.33	98.53	64.10
	NOTCH2	**0.02**	4.73	5.30	4.08	94.12	56.41
	PIGR	**<0.001**	1.51	3.23	67.35	60.29	63.25
	PPP1CA	**<0.001**	5.11	6.34	2.04	98.53	58.12
	ROCK2	**0.99**	6.81	5.37	0.00	95.59	55.56
	SDF2L1	**<0.001**	3.00	5.03	18.37	89.71	59.83
	TOPORS	**<0.001**	1.54	3.99	75.51	63.24	68.38

Table 4.4 Differential expression evaluated by qRT-PCR between Pt-r and Pt-s samples of the 23 selected genes in cohort B. Average expression of Pt-s and Pt-r samples and t-test p-value is reported. In cohort B, Pred Pt-s/Pt-r (%) is calculated based on the threshold of cohort A, and the accuracy is the total number of patients classified correctly. In gray are reported those genes that are not validated. Table is divided in up and down regulated genes in order to have odd ratios always greater than 1: for genes up-regulated in Pt-r patients the reference group is Pt-r, for genes up-regulated in Pt-s the reference group is Pt-s.

In order to quantify the prediction power of these 17 genes, we estimated the expression threshold (defined as the expression cut-off maximizing the sum of specificity and sensitivity) for each gene in cohort A and evaluated the accuracy of the prediction in cohort B using these thresholds. As expected, in cohort A all the 17 genes showed excellent levels of specificity (min 75% - max 100%) and sensitivity (min 75% - max 95%), while in cohort B we observed high levels of accuracy, defined as correctly predicted classes on the total number of testing classes, (>70%) for *CACNA1C*, *FLNC*, *PPPR12A*, *PRKG1*, *APOL6* genes. The list of 17 validated genes is referred from now onwards to as *SI* signature.

4.1.4 Prognostic performance of *SI* signature

The presence of an association between signature *SI* and the intrinsic resistance to Pt-based chemotherapy does not directly implicate that signature *SI* is also associated with survival parameters. Thus, to verify whether any of the elements of signature *SI* could have an additional prognostic role, qRT-PCR expression values have been associated with prognostic parameters (PFS and OS) both in cohort A and B. To increase the sample size, analysis was performed including also Pt-partially sensitive (Pt-ps) patients enrolled in cohort A (n=19) and in cohort B (n=25). To test the association between signature *SI* and OS a total number of 242 patients derived from cohort A and B were considered. Since in cohort B PFS follow up information are missing for four patients, the association between signature *SI* and PFS was investigated on 138 cases. The univariate and the multivariate Cox proportional hazards models adjusted for age and RT were used to access the prognostic significance of each gene. As reported in table 4.5, 14 out of the 17 genes (82%) composing signature *SI* are associated with both PFS and OS in uni- and in multivariate analysis.

	OS								PFS							
	UNIVARIATE				MULTIVARIATE				UNIVARIATE				MULTIVARIATE			
GENE	HR	CI 95% lower	CI 95% upper	P	HR	CI 95% lower	CI 95% upper	P	HR	CI 95% lower	CI 95% upper	P	HR	CI 95% lower	CI 95% upper	P
APOL6	0.52	0.38	0.71	<0.001	0.54	0.39	0.74	<0.001	0.44	0.32	0.60	<0.001	0.47	0.34	0.63	<0.001
BAD1	0.49	0.36	0.67	<0.001	0.49	0.36	0.67	<0.001	0.61	0.46	0.81	<0.001	0.57	0.43	0.76	<0.001
CACNA1C	2.01	1.46	2.77	<0.001	2.08	1.50	2.88	<0.001	2.05	1.53	2.75	<0.001	2.23	1.65	3.00	<0.001
CREB3	0.50	0.36	0.68	<0.001	0.53	0.39	0.74	<0.001	0.50	0.37	0.68	<0.001	0.51	0.38	0.70	<0.001
EYA4	1.84	1.17	2.88	0.01	1.79	1.14	2.81	0.01	1.73	1.09	2.74	0.02	1.74	1.09	2.76	0.02
FANCA	0.54	0.39	0.75	<0.001	0.53	0.38	0.73	<0.001	0.55	0.41	0.74	<0.001	0.53	0.40	0.72	<0.001
FANCI	0.45	0.32	0.64	<0.001	0.51	0.36	0.73	<0.001	0.56	0.41	0.76	<0.001	0.60	0.44	0.82	<0.001
FLNC	2.45	1.77	3.38	<0.001	2.49	1.80	3.45	<0.001	2.77	2.02	3.80	<0.001	2.93	2.12	4.04	<0.001
HLA-A	0.60	0.42	0.86	0.01	0.71	0.49	1.02	0.07	0.53	0.37	0.76	<0.001	0.62	0.43	0.89	0.01
HLA-F	0.45	0.33	0.61	<0.001	0.50	0.37	0.69	<0.001	0.44	0.33	0.59	<0.001	0.47	0.35	0.62	<0.001
NOTCH2	0.62	0.40	0.96	0.03	0.65	0.42	1.01	0.06	0.65	0.43	0.97	0.04	0.65	0.43	0.98	0.04
PIGR	0.68	0.49	0.93	0.02	0.61	0.44	0.84	<0.001	0.65	0.49	0.88	<0.001	0.58	0.43	0.79	<0.001
PRKG1	4.13	2.75	6.21	<0.001	3.52	2.31	5.36	<0.001	6.66	4.38	10.13	<0.001	5.66	3.69	8.66	<0.001
STAT1	0.83	0.60	1.14	0.24	0.99	0.72	1.36	0.94	0.84	0.60	1.16	0.29	0.98	0.70	1.37	0.90
TOPORS	0.51	0.36	0.72	<0.001	0.49	0.34	0.70	<0.001	0.48	0.35	0.67	<0.001	0.46	0.33	0.63	<0.001
PPP1R12A	2.59	1.82	3.70	<0.001	2.69	1.88	3.85	<0.001	3.04	2.17	4.26	<0.001	3.32	2.36	4.67	<0.001
SDF2L1	0.47	0.34	0.65	<0.001	0.49	0.35	0.67	<0.001	0.43	0.32	0.58	<0.001	0.43	0.32	0.58	<0.001

Table 4.5 Univariate and multivariate survival models of the 17 DEG elements enrolled within the *SI* signature. For each gene the hazard ratio (HR) with 95% confidence interval and p-value (P) is reported for either OS and PFS. In grey are highlighted the three genes not associated with level of significance in uni and multivariate analysis to both OS and PFS.

PIGR shows a non-significant correlation in univariate analysis with OS (p-value = 0.02). However, as PIGR have demonstrate a very strong association with OS in multivariate analysis (p-value < 0.001), it has been considered to have prognostic significance. The 14 genes with prognostic relevance identified have been defined as signature *SII.*

4.1.5 Signature *SII* validation across an external dataset

To further confirm the prognostic relevance of signature *SII* we used gene expression profiles of an external cohort of HGS-EOC patients (cohort C, Curated Ovarian Cancer Dataset). From the entire Curated Ovarian Cancer Database (n=2970), 838 Pt-treated patients with complete follow-up including PFS and OS information have been selected for this study.

As anticipated in paragraph 4.1.1, samples selected from cohort C do not have PFI information required for the correct Pt-s and Pt-r classification. As a proxy of PFI, PFS has been used to classify patients into Pt-s (with PFS>12 months) and Pt-r (with PFS<6 months). Although this approach is clearly biased, it will give us an indication of the prognostic potential of our signature. By comparing the expression data available for the 838 cases, we observed that the expression profiles of ten out of 14 genes of *SII* signature (71%) resulted as differentially expressed between Pt-s and Pt-r cases (Table 4.6). Finally, three out these ten genes, namely *PRKG1, SDF2L1 and PPP1R12A* we significantly associated with level of significance in both uni- and multi-variate, adjusted for age and RT, analyses with OS and PFS.

	Pt-r vs Pt-s	OS								PFS UNIVARIATE							
		UNIVARIATE				MULTIIVARIATE				UNIVARIATE				MULTIIVARIATE			
GENE	Combined p	HR	CI 95% lower	CI 95% upper	P	HR	CI 95% lower	CI 95% upper	P	HR	CI 95% lower	CI 95% upper	P	HR	CI 95% lower	CI 95% upper	P
PRKG1	0.0006	1.18	1.08	1.30	**<0.001**	1.15	1.05	1.27	**<0.001**	1.14	1.05	1.24	**<0.001**	1.14	1.05	1.24	**<0.001**
SDF2L1	0,044751	0.84	0.77	0.92	**<0.001**	0.85	0.78	0.93	**<0.001**	0.87	0.8	0.94	**<0.001**	0.88	0.82	0.96	**<0.001**
PPP1R12A	0,011071	1.15	1.05	1.26	**<0.001**	1.12	1.02	1.23	**0.01**	1.11	1.02	1.2	**0.01**	1.10	1.02	1.19	**0.02**
FANCI	0,001197	0.98	0.90	1.07	**0.70**	1.00	0.92	1.09	**0.98**	0.94	0.87	1.02	**0.12**	0.93	0.86	1.01	**0.07**
TOPORS	0,170422	0.90	0.83	0.98	**0.02**	0.93	0.85	1.01	**0.10**	0.93	0.86	1.01	**0.09**	0.93	0.86	1.01	**0.07**
CREB3	0,136748	0.87	0.79	0.95	**<0.001**	0.89	0.81	0.97	**0.01**	0.9	0.84	0.98	**0.01**	0.93	0.86	1.10	**0.08**
FLNC	0.016	1.11	1.02	1.21	**0.02**	1.10	1.01	1.20	**0.04**	1.07	0.99	1.15	**0.08**	1.06	0.98	1.15	**0.12**
FANCA	0,003142	0.99	0.90	1.07	**0.73**	1.01	0.92	1.09	**0.90**	0.95	0.88	1.03	**0.23**	0.95	0.88	1.03	**0.26**
EYA4	6,74E-06	1.13	1.04	1.23	**<0.001**	1.08	0.99	1.18	**0.09**	1.06	0.98	1.14	**0.16**	1.04	0.96	1.13	**0.30**
PIGR	0,030124	0.92	0.84	1.02	**0.10**	0.95	0.87	1.05	**0.32**	1	0.93	1.08	**0.93**	1.03	0.96	1.11	**0.46**
HLA-F	0,023082	0.95	0.87	1.03	**0.24**	0.99	0.90	1.07	**0.75**	0.95	0.88	1.03	**0.24**	0.97	0.90	1.05	**0.48**
CACNA1C	0,051102	1.02	0.94	1.12	**0.59**	1.01	0.93	1.11	**0.76**	1.02	0.95	1.11	**0.55**	1.02	0.94	1.10	**0.69**
APOL6	0,011968	0.92	0.85	1.00	**0.05**	0.96	0.88	1.05	**0.38**	0.99	0.92	1.07	**0.78**	1.01	0.94	1.09	**0.78**
BAD1	0,216293	1.01	0.93	1.10	**0.86**	1.00	0.92	1.10	**0.94**	0.99	0.92	1.07	**0.82**	1.01	0.93	1.09	**0.88**

Table 4.6 Univariate and multivariate survival models of *SII* signature in cohort C. In the first column the Tippet's combined p-value of Youen t-test between Pt-s and Pt-r samples is reported. For each gene the HR value with 95% confidence interval and p-value (P) is reported for either OS and PFS. In bold is reported the list of ten genes the expression level of which is different between Pt-s and Pt-s cases in cohort C. In yellow are highlighted the three genes (*PRKG1, SDFL1* and *PPP1R12A*) associated with level of significance in both uni and multivariate analysis to both OS and PFS. These three genes constitute the *SIII* signature.

Patients with high expression levels of *PRKG1* and *PPP1R12A* and low expression levels of *SDF2L1* had a shorter survival than those with low levels of *PRKG1* and *PPP1R12A* and high levels of *SDF2L1* (Table 4.6).

These three genes are referred from now onwards to as *SIII* signature.

4.1.6 Signature combination

In order to test a possible synergistic effect among the three genes composing the signature *SIII*, patients have been stratified into three classes:

- high risk: high levels of PPPR12A and PRKG1 and low levels of SDF2L1
- low risk: low levels of PPPR12A and PRKG1 and high levels of SDF2L1
- intermediate risk: all the other combinations of expression values

On the basis of this classification the ability of the combination of three elements of the signature *SIII* have been tested (figure 4.3). KM curves demonstrated significant differences in terms of PFS and OS between patients with high (red line) and low risk (green line) both in cohort A and in cohort B. Patients characterized by an intermediate risk (blue line) are, as expected, located between the other two groups.

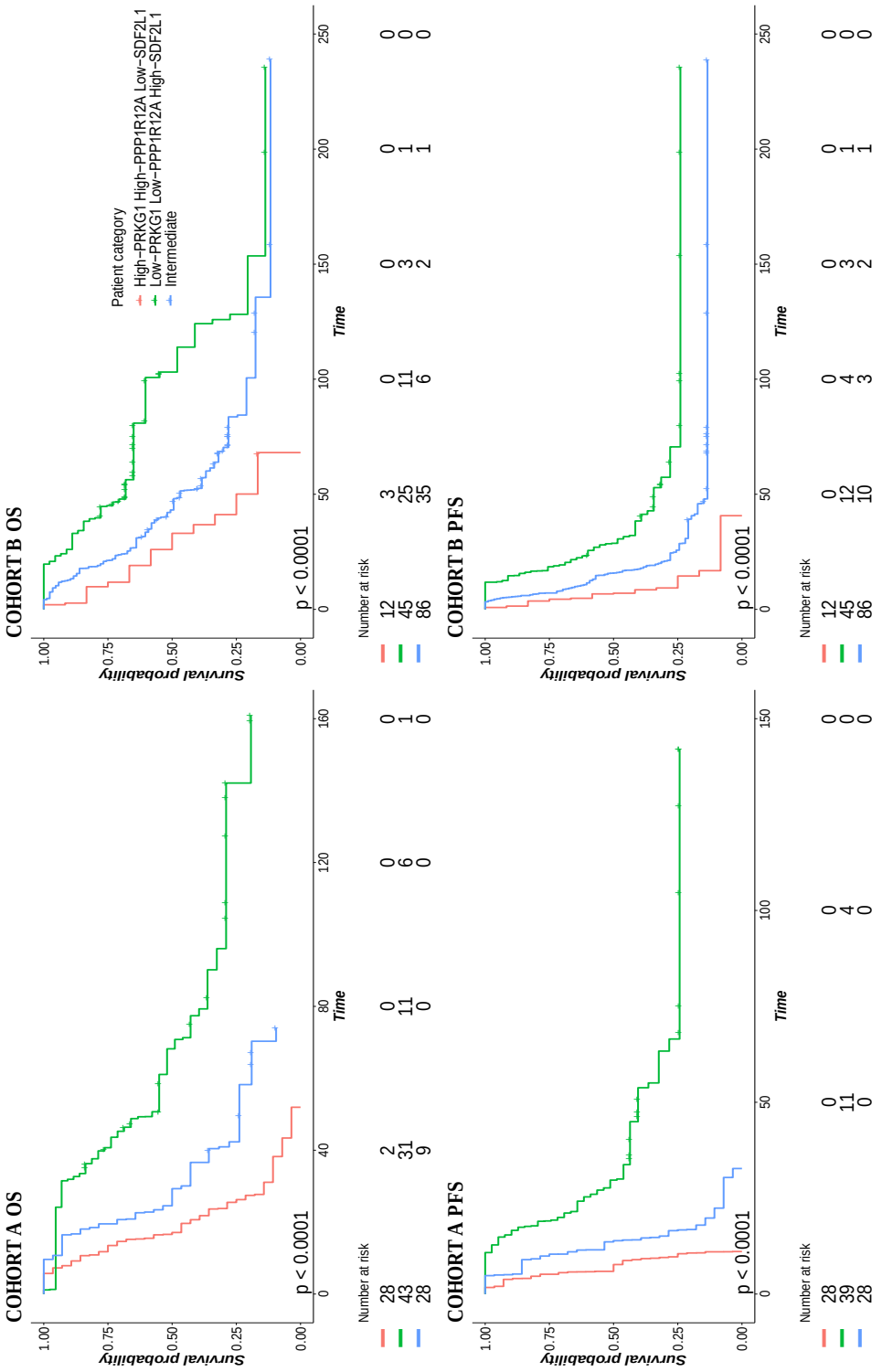


Figure 4.3 Kaplan Mayer curves of the combination of the expression values of PRKG1, SDF2L1 and PPPR12A. Kaplan Mayer curves and log-rank test *p*-values of the combination of the expression values of *PRKG1*, *SDF2L1* and *PPPR12A* genes either using OS (panel A) and PFS (Panel B) as survival measures.

As the most important prognostic factor in HGS-EOC is represented by RT after surgery, the robustness of signature *SIII* in patients' prognosis has been also verify considering this parameter. KM curves reported in figure 4.4 demonstrate that the significant stratification based

on signature *SIII* is maintained even when the RT status is considered.

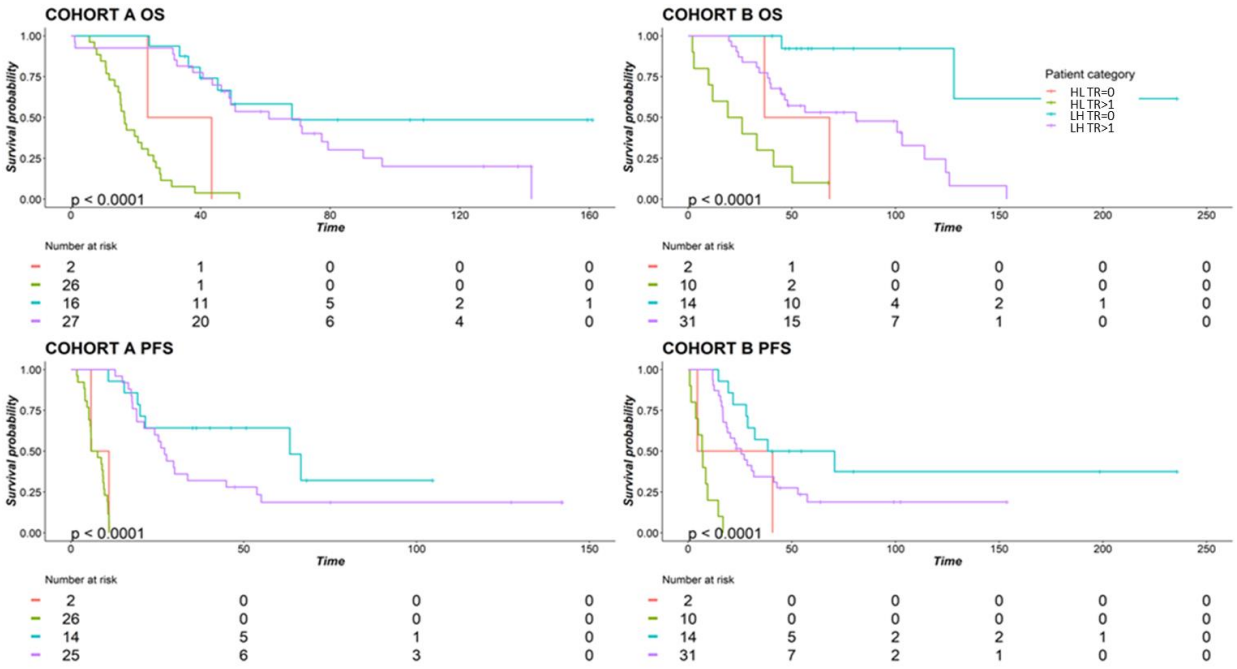


Figure 4.4 Kaplan Mayer curves of the combination of the expression values of *PRKG1*, *SDF2L1* and *PPPR12A* with stratification by patient TR status. Kaplan Mayer curves and log-rank test p-values of the combination of the expression values of *PRKG1*, *SDF2L1* and *PPPR12A* genes either using OS and PFS as survival measures, with stratification by patient TR status. (HL = "High-PRKG1 High-PPP1R12A Low-SDF2L1"; LH = "Low-PRKG1 Low-PPP1R12A High-SDF2L1")

This result confirms that the combination of expression level of genes belonging to the *SIII* signature is able to significantly stratify patients with different prognosis both in OS and PFS.

4.1.7 Conclusion

The goal of this first part of the work was to identify a regulatory network built up by annotated coding genes and non-coding genes able to explain part of the complex biological mechanisms underlying the intrinsic Pt resistance in HGS-EOC patients. To this aim, a network composed by 131 genes and five miRNAs, that belong to five different and interconnected biological processes, has been identified. Three genes of this network, including the hub represented by

PRKG1 gene, had been demonstrated both predictive of Pt response and prognostic in OS and PFS, suggesting their possible role as prognostic biomarkers associated with survival.

4.2 Analysis of genomic instability

Since the data reported in the literature demonstrate associations between alterations in DNA repair mechanisms (i.e. Homologous recombination, HR) and Pt response (80) (132) (84), in this second part of this result section mutational status of *BRCA* genes and HR pathway were evaluated in a sub-cohorts of our HGS-EOC patients (AI=28 and BI=55). In particular, studying the "genomic scars" due to HR deficiency (see Introduction section, paragraph 1.4.2.2), such as large-scale transition (LST), loss of heterozygosis (LOH) and telomeric allelic imbalances (TAI). An in-house protocol for HRD score was calculated and a threshold set and tested in order to verify its ability to discriminate Pt-s from Pt-r cases. The association of HR deficiency/proficiency status calculated on the basis of our HRD score with Pt response could have independent value with other predictive biomarkers in stratifying HGS-EOC patients on the basis of their Pt response at time of diagnosis.

4.2.1 AI and BI Cohorts' Description

Cohorts AI and BI are derived from previously selected cohorts A and B (paragraph 4.1.1). The median age at diagnosis was 66 and 54 years for cohort AI and BI, respectively (table 4.1), and the median follow-up time was respectively two and four years for cohort A and cohort BI, respectively. Being the clinical parameters comparable with those reported in literature (Curated Ovarian dataset, table 4.1) it was possible to assume that our cohorts well represented the clinical features of HGS-EOC. In our cohorts, the Progression Free Survival (PFS) was 10.9 and 20.4 months and the Overall Survival (OS) was 31.8 and 48.5 months in cohort AI and BI, respectively. The majority of patients in both cohorts showed a suboptimal residual tumour (RT) after surgery (64% in Cohort AI and 69% in Cohort BI), that means that the tumour mass still remaining after surgery is > 1cm (R2). Moreover, as the study is based on the identification of molecular signatures associated to Pt response, patients of the two cohorts were best balanced in terms of platinum sensitivity. In particular, 14 and 30 Pt-sensitive and 14 and 25 Pt-resistant patients were selected in Cohort AI and BI respectively. HGS-EOC patients that were refractory or partially sensitive to chemotherapy were excluded.

Clinico-pathological features of HGS-EOC (FIGO stage III/IV) patients composing Cohort AI (n=28) and Cohort BI (n=55), are summarised in table 3.7. Although there are differences in terms of PFS and OS between the two cohorts, likely due to the low number of cases and the specific patients selection (extremely sensitive and extremely resistant patients), they are not statistically significant (p=0.25 and p=0.27 respectively, test Mann-Whitney).

4.2.2 Characterization of HGS-EOC snap-frozen biopsies

The starting material exploited for the biological experiments reported in this section of the thesis is represented by nucleic acids extracted and purified from HGS-EOC snap-frozen biopsies. These samples were analysed for:

i) Tumour purity

ii) Ploidy

iii) TP53 clonal mutation or MDM2/MDM4 amplifications

Tumour purity and the ploidy were investigated *in silico* both to have an indication of tumour continent of biopsies analysed and to correct the results obtained from mutational and CNV analysis. Moreover, the *TP53* clonal pathogenic (IARC *TP53* database) mutations and the *MDM2/MDM4* amplifications (log-ratio >|0.1|), were respectively evaluated as they represent two clonal events that characterize HGS-EOC disease, providing a genetic confirmation of the histological tumour classification.

Table 4.7 reports for each case the above mentioned parameters.

Cohort	ID	Purity	Ploidy	*TP53*	*MDM4*	*MDM2*
AI	SC-1	80%	1,82	-	AMP	AMP
	SC-12	63%	1,92	c.712T>C (44.46%); LOH	-	AMP
	SC-13	54%	1,92	-	AMP	DEL
	SC-15	88%	1,94	c.2077G>A (9.37%); LOH	-	-
	SC-16	68%	1,91	c.159G>A (66.63%); LOH	-	-
	SC-22	73%	4,13	c.401T>G (80.45%); LOH	AMP	AMP
	SC-23	88%	3,11	-	AMP	AMP
	SC-25	38%	3,52	c.1013_1014insA (33.93%); LOH	AMP	AMP
	SC-31	95%	1,74	c.524G>A (95.20%); LOH	-	-
	SC-32	56%	3,13	-	AMP	AMP
	SC-40	83%	1,84	-	AMP	-
	SC-43	81%	2,87	c.743G>A (80.99%); LOH	AMP	AMP
	SC-50	67%	3,12	c.659A>G (73.29%); LOH		
	SC-51	73%	1,88	c.993+1G>A (51.08%); LOH	-	AMP

	SC-52	73%	2,99	-	AMP	AMP
	SC-54	56%	3,07	c.524G>A (63.73%); LOH	AMP	AMP
	SC-56	58%	4,59	c.646G>A (84.05%); LOH	AMP	AMP
	SC-58	45%	1,75	c.661G>T (43.93%); LOH	-	-
	SC-62	21%	1,85	c.736A>G (8.55%); LOH	-	-
	SC-67	40%	3,40	c.548C>G (38.13%); LOH	AMP	AMP
	SC-72	28%	1,80	c.526T>A (50%)	-	-
	SC-74	66%	3,77	c.527G>T (76.93%); LOH	AMP	AMP
	SC-77	50%	3,56	-	AMP	AMP
	SC-87	81%	2,57	c.734G>T (75.75%); LOH	AMP	AMP
	SC-96	28%	2,69	c.579dup (27.12%); LOH	-	AMP
	SC-97	18%	2,05	-	-	AMP
	SC-98	31%	2,49	c.714dup (26.50%); LOH	AMP	AMP
	SC-99	78%	1,75	c.319del (62.26%); LOH	-	-
BI	**10023**	61%	1,77	c.578A>G (46.14%); LOH	AMP	-
	10104	92%	3,96	-	AMP	AMP
	20006	55%	2,33	c.469G>T (57.99%); LOH	AMP	AMP
	20041	79%	3,17	c.517G>A (78.96%); LOH	AMP	AMP
	20132	81%	2,98	-	AMP	-
	20155	55%	3,00	-	AMP	AMP
	20614	17%	1,78	-	AMP	-
	20768	81%	2,55	c.722C>T (90.78%); LOH	AMP	-
	20777	90%	2,05	-	AMP	-
	20789	44%	3,18	c.707A>G (51.83%); LOH	-	AMP
	20828	64%	1,91	-	AMP	DEL
	20829	48%	3,40	-	AMP	AMP
	20830	60%	1,99	c.949C>T (55.26%); LOH	-	-
	20935	48%	3,32	-	AMP	AMP
	20937	90%	2,78	c.713G>T (92.07%); LOH	AMP	AMP
	20956	42%	3,90	c.653T>G (47.34%); LOH	AMP	AMP
	20974	41%	2,90	c.432_446del (32.80%); LOH	AMP	AMP
	20976	94%	5,08	c.488A>G (78.53%); LOH	AMP	AMP
	20979	30%	1,73	c.713G>T (18.78%); LOH	-	-
	20983	15%	2,02	c.280dup (5.22%); LOH	-	-
	21021	69%	1,96	c.742C>T (48.76%); LOH	-	-
	21047	71%	1,88	c.584T>C (56.15%); LOH	-	-
	21089	22%	2,31	-	AMP	-
	21095	95%	2,70	c.637C>T (95.30%); LOH	AMP	AMP
	21096	50%	2,99	c.743G>A (43.58%); LOH	AMP	AMP
	21113	78%	4,07	c.733G>A (82.09%); LOH	AMP	AMP
	21117	72%	3,43	c.686_687del (78.89%); LOH	AMP	AMP
	21118	71%	2,68	c.637C>T (75.00%); LOH	AMP	AMP
	21121	95%	5,47	-	AMP	AMP
	21128	18%	1,96	c.1116A>C (67%)	-	-
	21140	63%	3,63	c.586C>T (66.77%); LOH	AMP	AMP
	21141	15%	2,07	c.529_546del (62%)	-	-
	21150	76%	3,47	-	AMP	AMP
	21151	65%	2,64	c.536A>T (67.90%); LOH	AMP	AMP
	21157	40%	2,99	-	-	AMP
	21159	48%	1,90	c.637C>T (34.61%); LOH	-	-

	21171	75%	1,86	c.437G>A (57.40%)	-	AMP
	21175	18%	2,02	c.1014C>T (92.3%)	-	-
	21189	48%	3,98	c.162_165dup (59.20%); LOH	AMP	AMP
	21193	23%	1,89	c.919+1G>T (16.39%); LOH	-	-
	21197	28%	2,11	c.240del (25.37%); LOH	AMP	-
	21232	49%	2,90	c.524G>A (44.99%); LOH	AMP	AMP
	21238	64%	1,82	c.568C>A (48.85%); LOH	-	AMP
	21239	36%	1,87	-	AMP	AMP
	21251	59%	2,18	c.527G>A (66.76%); LOH	-	-
	21272	55%	2,90	c.584T>C (48.08%); LOH	AMP	AMP
	21288	55%	2,56	c.659A>G (39.89%); LOH	AMP	AMP
	21291	44%	1,67	c.1024C>T (49.35%); LOH	-	DEL
	21296	95%	3,58	c.775G>T (46.08%); LOH	AMP	AMP
	21311	47%	1,88	c.638G>T (37.19%); LOH	AMP	-
	21329	82%	1,86	c.919+1del (77.42%); LOH	-	DEL
	21331	20%	2,14	c.672+1G>T (14.50%); LOH	AMP	-
	21343	70%	1,98	c.924G>T (45.7%)	-	-
	21350	55%	2,74	c.743G>A (45.25%); LOH	AMP	-
	21358	34%	2,75	-	AMP	AMP

Table 4.7 *TP53* clonal mutations and *MDM4/MDM2* amplification status in cohort AI and BI. In table are reported the sample ID and the cohort to which they belong. For each sample the tumour percentage (purity) and the ploidy value are reported. In column *TP53* is reported the clonal mutation identified for each samples, the allelic fraction (%) and, where reported, the loss of heterozygosis (LOH), *MDM4/MDM2* copy number status is also reported (AMP = amplification, DEL = deletion). "-" Indicates that neither clonal *TP53* mutation nor *MDM4/MDM2* amplification/deletion were identified.

The tumour purity of the samples selected ranges from 18% to 95% and from 15% to 95% in cohorts AI and BI respectively, revealing a great variability in tumour content of biopsies analysed. Considering the ploidy as an important parameter that need to be taken into account in particular for CNV calculation, results obtained indicated that ploidy values ranging from 1,7 to 4,6 in cohort AI and from 1,7 to 5,5 in cohort BI.

The clonal pathogenic *TP53* was identified in 20 (71%) and in 41 (74%) cases of cohorts AI and BI respectively, whereas MDM2/MDM4 amplifications in all the other samples composing the two cohorts (table 4.7). Considering both cohorts, the allelic fraction (AF) of *TP53* mutations was variable, according to the tumour purity of sample analysed.

Taking together these results confirm the impairment of the *TP53* function as the main molecular feature of HGS-EOC

4.2.3 Mutational status of *BRCA1* and *BRCA2* genes

As nowadays it is well established that tumours carrying somatic or germline pathogenic mutations in tumour suppressor *BRCA* genes are more responsive to DNA-damaging agents (i.e. Pt compounds), the mutational status of *BRCA1* (Ch13) and *BRCA2* (Ch17) was investigated (Table 4.8)

Cohort	ID	*BRCA* mutational status			Pt-status
		g*BRCA*	*BRCA1*m	*BRCA2*m	
AI	SC-1				Res
	SC-12				Sens
	SC-13	*BRCA2*			Sens
	SC-15		c.2077G>A (9.37%); LOH		Sens
	SC-16	*BRCA2*			Sens
	SC-22				Res
	SC-23				Sens
	SC-25				Res
	SC-31				Sens
	SC-32				Res
	SC-40			c.2971A>G (15.00%); LOH	Sens
	SC-43	*BRCA2*			Sens
	SC-50		c.3748G>T (88.56%); LOH		Sens
	SC-51				Sens
	SC-52				Res
	SC-54		c.1067A>G (16.58%); LOH		Sens
	SC-56				Sens
	SC-58		c.649del (24.56%); LOH		Res
	SC-62				Res
	SC-67				Res
	SC-72	*BRCA2*			Res
	SC-74		c.2630del (68.57%); LOH		Sens
	SC-77				Res
	SC-87				Res
	SC-96				Res
	SC-97				Res
	SC-98				Res
	SC-99				Sens
BI	10023		c.2008G>T (42.98%); LOH		Sens
	10104				Res
	20006			c.1114A>C (99.04%); LOH	Res
	20041				Res
	20132				Sens
	20155		c.2612C>T (77.49%); LOH	c.1114A>C (15.90%); LOH	Sens

	20614				Sens
	20768				Res
	20777				Sens
	20789	*BRCA1*			Res
	20828				Res
	20829				Sens
	20830	*BRCA1*			Sens
	20935				Sens
	20937				Res
	20956				Sens
	20974				Sens
	20976				Sens
	20979				Res
	20983				Sens
	21021				Res
	21047				Sens
	21089				Res
	21095		c.1399A>T (100.00%); LOH		Sens
	21096				Sens
	21113				Res
	21117				Sens
	21118				Res
	21121				Res
	21128				Sens
	21140	*BRCA1*			Sens
	21141			c.6761_6762del (51.69%); LOH	Sens
	21150				Sens
	21151				Res
	21157				Res
	21159				Sens
	21171				Res
	21175			c.1114A>C (5.70%); No LOH	Res
	21189				Sens
	21193				Sens
	21197				Res
	21232				Res
	21238	*BRCA1*			Sens
	21239				Res
	21251				Sens
	21272	*BRCA1*			Res
	21288				Sens
	21291				Res
	21296				Res
	21311	*BRCA1*			Sens
	21329				Sens
	21331				Res
	21343				Res

	21350				Sens
	21358				Res

Table 4.8 Mutational status of *BRCA1* and *BRCA2* genes in cohort AI and BI. For each case is specified the nature of the pathogenic (BRCA Excange) *BRCA* mutation, germline mutation (g*BRCA*) or somatic mutation (*BRCA1*m and *BRCA2*m). The allelic fraction (AF) for *BRCA*m mutations are reported in paremtheses. Platinum status sensitivity is reported for each case (Res, resistant and Sens, sensitive). For cases highlighted in blue no matched blood sample was available.

For those cases for whom the blood or normal matched tissue was available, pathogenic (BRCA Exchange) germline mutations in BRCA genes (g*BRCA*) were evaluated. For those cases with no normal reference DNA (five cases for cohort AI and seven patients for BI) the presence of g*BRCA* mutations could be only speculated on the basis of high allelic fractions of BRCA somatic mutations obtained in their tumour biopsies.

Considering the importance of these two *BRCA* genes in therapy response, somatic mutations were also considered. To identify *BRCA* mut in tumour biopsies, arbitrary stringent parameters were established:

- at least four mutated reads out of the total number of reads that passed the quality control
- a minimum coverage of 50X
- AF>5%
- Genetic variants reported as “pathogenic” or “likely pathogenic” in BRCA Exchange database

On the bases of these criteria in cohort AI five somatic *BRCA1* mutations with AF ranging between 9.37% and 88.56%, and one in *BRCA2* genes (AF 15%) were identified. These six somatic *BRCA* mutations were predominantly identified in Pt-s patients (five out of six), whereas one (*BRCA1* c.649del (24.56%) refers to Pt-r case. In cohort BI three *BRCA1* (42.98% < AF < 100%) and four *BRCA2* (5.70%< AF < 99.04) somatic mutations were respectively identified. Five out of seven *BRCA*mut derived from biopsies of Pt-s patients and the remaining two were associated to Pt-r cases.

Data reported above, confirmed that genetic mutations in *BRCA1* and *BRCA2* genes were not able to explain alone the sensitivity to Pt therapy, in fact, as reported, three patients with g*BRCA* mutations are clinically classified as Pt-r cases. For this reason, moving from single gene to pathway approach, the investigation regarding the status and the functional activity of the biological pathway in which *BRCA1* and *BRCA2* are involved in was needed.

4.2.4 Developing an HRD score

In paragraph 4.2.3 the germline and somatic status of *BRCA1* and *BRCA2* genes were tested, however these two genes are a small part of a more complex DNA repair pathway, the Homologous Recombination (HR). As the association between defects in HR pathway and Pt sensitivity has been already reported, in our cohort of patients a HR score able to discriminate HR deficient (HRD) from HR proficient cases was calculated and the distribution of patients according to their HR status was evaluated in association with their Pt response.

Considering cohorts AI and BI together, HRD score was calculated on 70 out of 83 HGS-EOC patients. The remaining 13 samples were in fact excluded due to their tumour content, as tumour purity lower than 30% was not sufficient for HRD analysis.

In order to set a HRD score threshold, 30 high quality sequenced samples were selected, and subdivided in two groups: the first group (n=11) characterized by a certain deficiency in HR, determined by the presence of somatic or germline pathogenic biallelic *BRCA* mutations and the second group (n=19) defined as HR proficient, as no somatic pathogenic biallelic mutations were identified in HR-related genes analysed. The biallelic condition imposed for the mutations considered, is due to the fact that some genes i.e. *BRCA* could maintain their functional activity also with the presence of only wt allele, so the application of this criteria allowed to clearly discriminate HR proficient from HR deficient cases. For each sample HRD scored was calculated and plotted on the basis of the group to whom it belonged (figure 4.5).

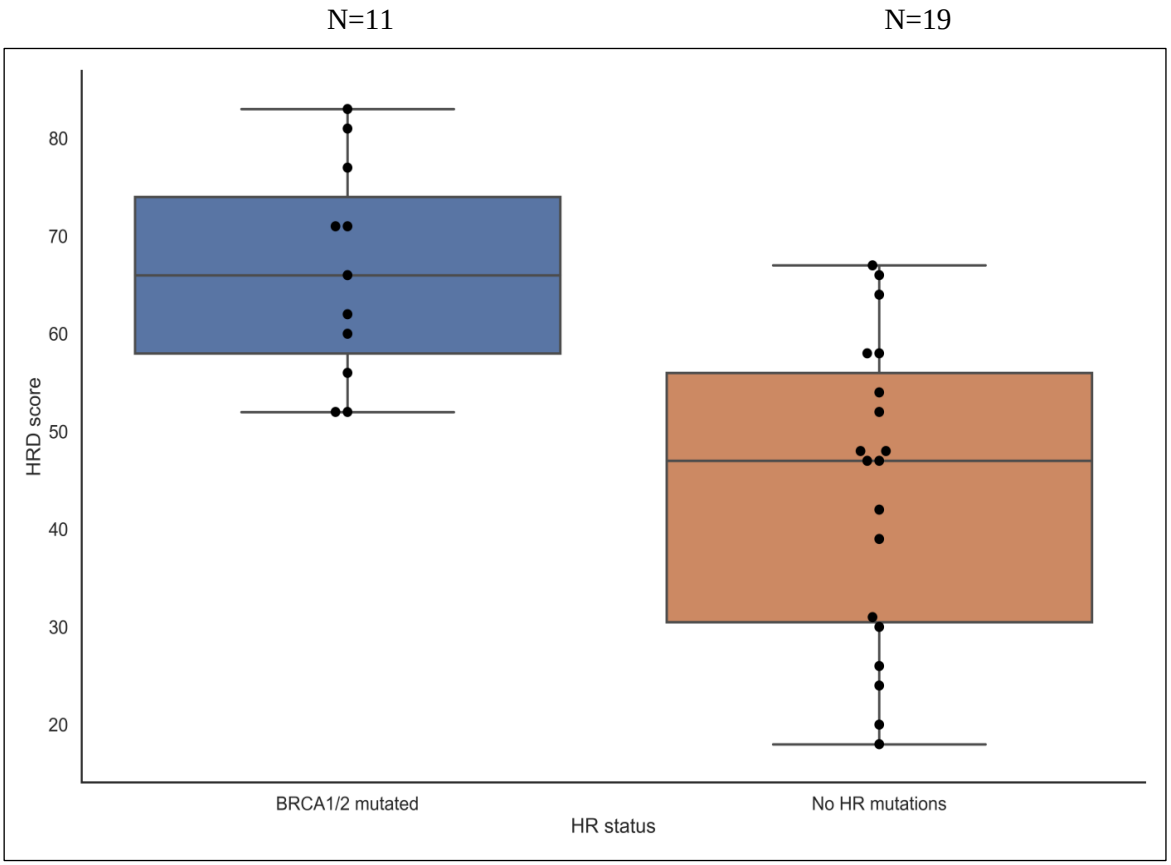


Figure 4.5 HRD Score in 30 high quality samples to set the cut-off. On the left (blue box-plot) the *BRCA1/2* mutated group and on the right (red box plot) samples without mutations in HR-related genes. For each group is reported the madian of value, the IRQ and the 5th and 95th percentile.

In figure 4.5 on the x-axis the *BRCA* mutated (blue blox plot) group and the wt HR-related genes (red box plot) are reported, whereas in the y-axis is reported the HRD score calculated. As expected, the median of HRD score was higher in *BRCA* mutated group versus the group without mutations in HR genes group (median 66 ranging from 52 to 83 versus median of 47 ranging from 18 to 67 respectively). In line with the threshold set in the gold standard HRD commercially available test (Myriad MyChoice) our cut-off defining HR deficient or HR proficient cases the was set at the 5th percentile of blue box plot (cut-off = 52).

To test whether the score threshold set was able di discriminate *BRCA* mutated samples from samples without mutations in HR genes, the other 40 cases not including in the first part of HRD analysis were used. For each sample HRD score was calculated and on the basis of the

cut-off decided, samples were subdivided into two groups, that are HRD score > 52 and HRD score < 52 (figure 4.6).

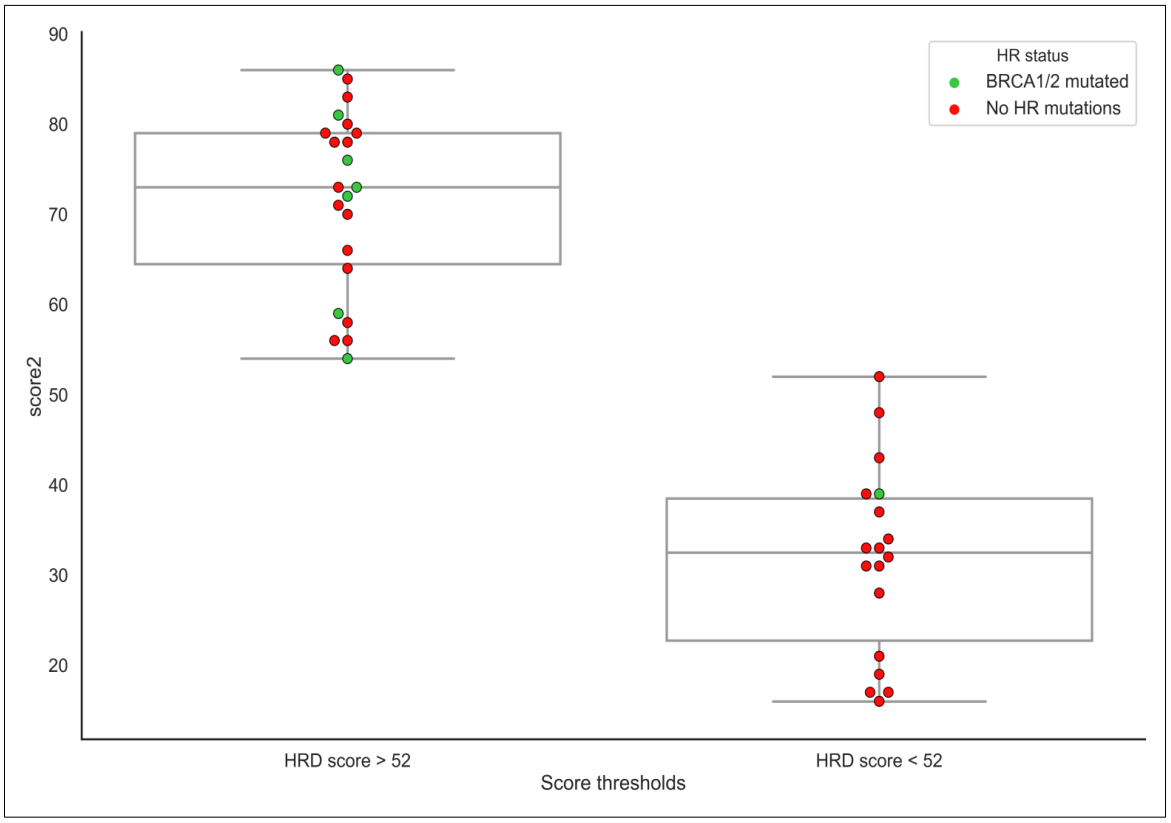


Figure 4.6 HRD score threshold test. On the y-axis the HRD score value is reported and on the x-axis the two groups of cases (HRD score > 52, on the left, and HRD score < 52, on the right) are reported. Red dots represent cases with BRCA mutations and green dots cases without mutations in HR-related genes.

The vast majority of *BRCA* mutated cases (green dots) was characterized by HRD score > 52 (87%, seven out of eight) whereas only one had HRD score < 52. Although, the distribution of wt HR genes-related samples (red dots) was not so defined (47% HRD score > 52 and 53% HRD score < 52) from a statistical point of view the distribution of samples into the two groups remained significant (p-value= 0.05), validating the HRD score threshold decided.

Basing on parameter defined (HRD score < or > 52), it was tested whether the deficiency or proficiency of HR was related to Pt sensitivity in the 70 cases selected for HRD analysis. To this aim the HRD score of Pt-s (n=40) and Pt-r (n=30) samples was calculated and plotted (figure 4.7).

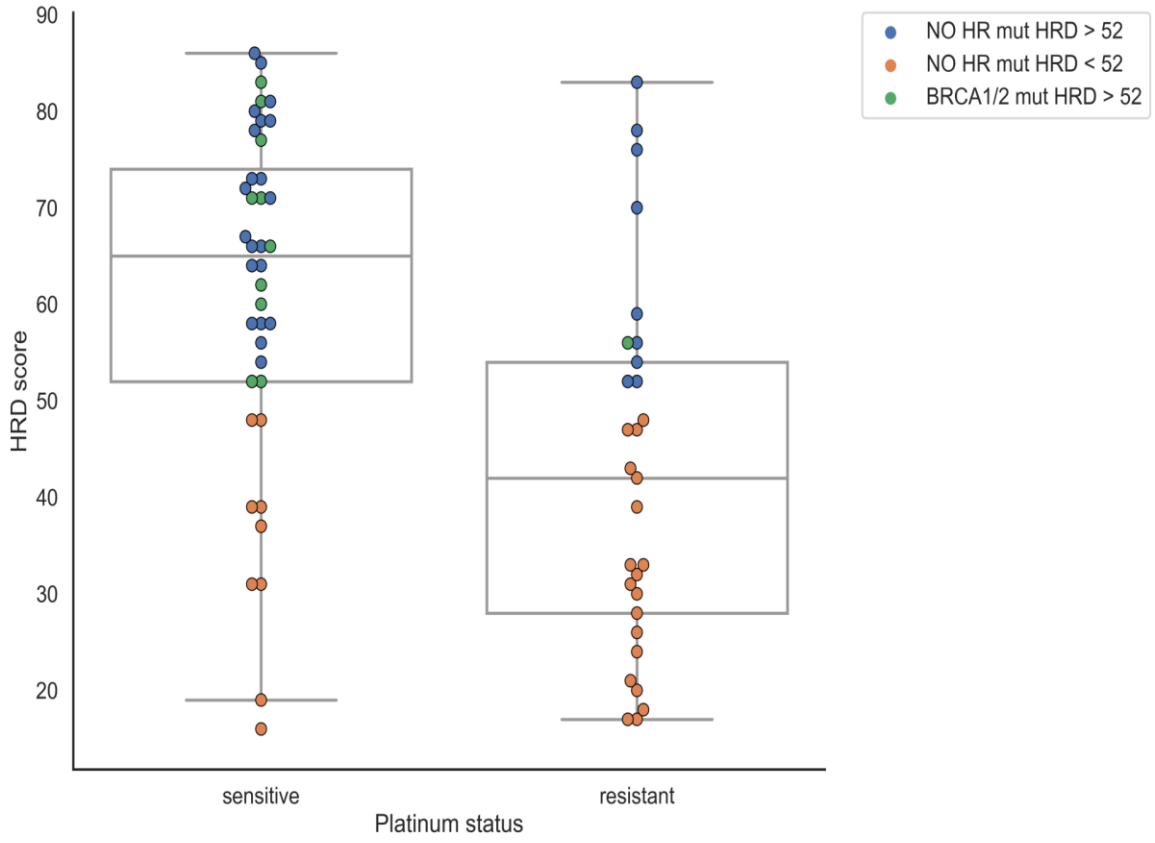


Figure 4.7 HRD score distribution in Pt-s (n=40) and Pt-r (n=30) cases. Median, IQR, 95th and 5th percentile are reported in the box plots. Each case (dot) is coloured on the basis of the figure legend reported according to its HRD score value and mutational status of HR-related and *BRCA* genes.

As shown in figure 3.8 Pt-s cases (box plot on the left) were mainly characterized by the presence of sample with HRD > 52 (31 out of 40), that are composed from both samples that carry BRCA mutations and from samples for whom no mutations in HR genes are reported. On contrary, Pt-r cases were enriched in samples with HRD score < 52 (20 out of 30).

This significant distribution (Fishers' test, p-value 0.0004) of sample with HRD > 52 in Pt-s cases, in which are included both *BRCA* mutated and wt HR-related genes, demonstrated that the evaluation of genomic consequences due to HR deficiency defined by HRD score >52, is able to better stratify patients on the basis of their Pt response rather then the evaluation of mutational status of *BRCA* genes alone.

4.2.5 Conclusion

In this second part of this work, the mutational status of *BRCA1/2* genes and the functional activity of the homologous recombination pathway, to whom they belong were evaluated. Although the academic in-house HRD-test that we are still developing, to date suffers of some limitations that are detailed described in paragraph 4.2, it had demonstrated to have a predictive value in HGS-EOC patients. However, data reported above demonstrated that the HRD score represents an important feature significantly associated with therapy sensitivity, but it is not able to explain alone the complex biological mechanisms behind the Pt response. For this reason, genetic information obtained by DNA sequencing needed to be completed by transcriptional information regarding both the known and the unknown part of the transcriptome. The integration of results obtained by the analysis of these two different aspects, could increase our capability in predicting patients' response to Pt therapy, and allow to better understand the biology related to Pt-resistance.

4.3 Transcriptomic analysis

The gene expression analysis reported in the first part of this results section (section 3.2), allowed to identify a regulatory network associated to intrinsic Pt resistance. However, as results reported were obtained exploiting microarray technique, the transcriptomic scenario analysed was limited to the annotated genes, not providing results related to i.e. the genome editing. The development and the diffusion of RNA sequencing technology over the past years allowed to complete the missing information and obtain a more detailed analysis of the transcriptome fine structure, overcoming limitations related to the microarray-based approaches.

To this aim, in this third part of this results section, the two sub-cohorts of HGS-EOC patients described in paragraph 4.2.1, named AI (n=28) and BI (n=55) were exploited. In particular, on cohort AI total RNA sequencing approach was used to identify differential expressed transcripts (DETs), both known and partially or total unknown, between Pt-r and Pt-s patients. To confirm results obtained a targeted RNA sequencing approach custom designed on DETs identified was then performed on both cohort AI (technical validation) and cohort BI (clinical validation).

4.3.1 Identification of differential expressed transcripts in Pt-r versus Pt-s patients

Total RNA sequencing experiments were performed on Cohort AI (see Results, paragraph 4.2.1). On the basis of the differential expression analysis in Pt-r versus Pt-s cases and with a FDR ≤ 0.05, a total of 1730 transcripts were identified as differentially expressed of which 1052 DETs resulted up-regulated ($2.11 < \text{Log FC} < 10.26$) and 678 down-regulated ($-8.72 < \text{Log FC} < -2.12$) (See appendix Table 6.1).

In order to better characterize the DETs identified, each transcript was reconstructed, compared to a reference annotation (Ensembl) and classified on the basis of the class codes reported in Gff Compare open-source software (Figure 4.8)

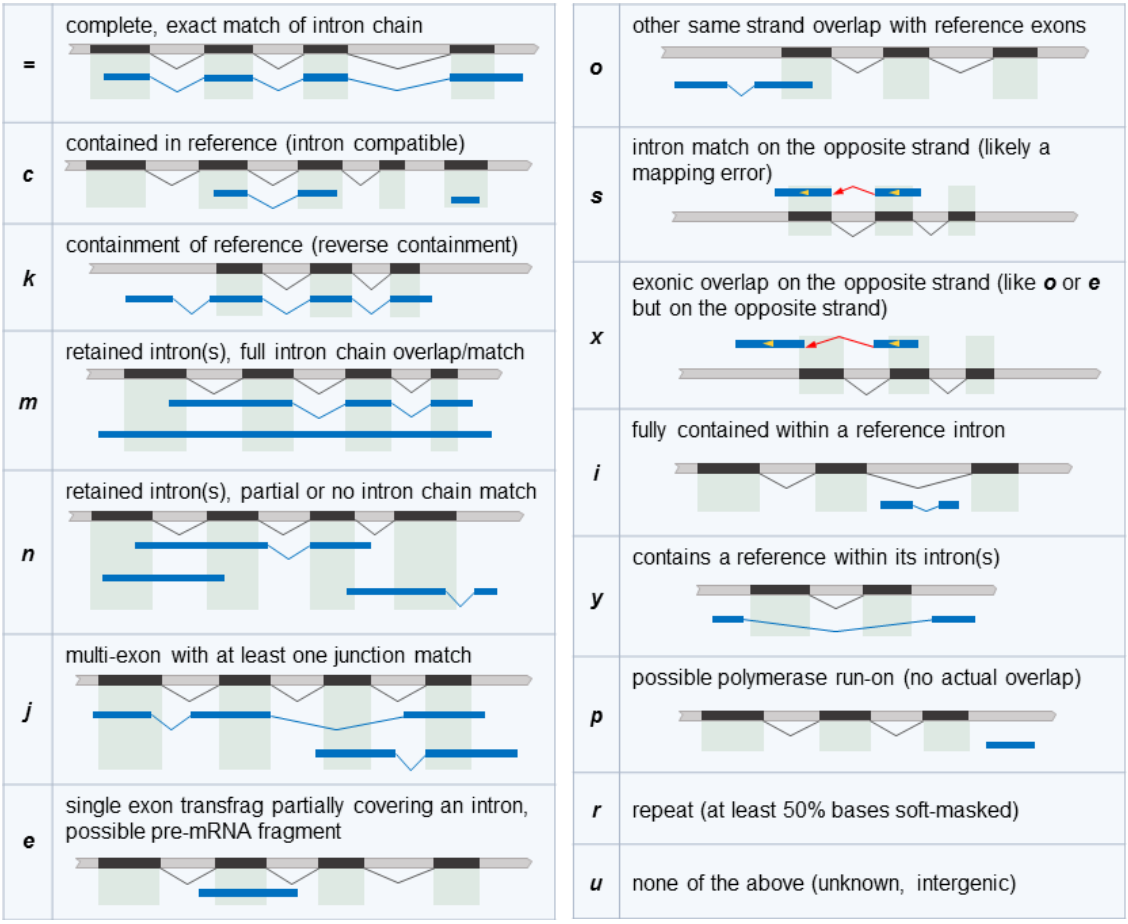


Figure 4.8 Gff Compare's class codes. In figure are reported the 15 Gff compare's class codes (=, c, k, m, n, j, e, o, s, x, i, y, p, r and u) that define the relationship between transcripts identified by RNA sequencing experiments and the closest reference transcript (where applicable). The class codes are shown in order of reference transcript similarity, form the complete match with reference transcript (class code =) to a completely unknown transcript which can not be associated with a reference sequence (class code "u")

In figure 4.9 is graphically represented the percentage of DETs for each class code. In particular, 562 out of 1730 DETs (32%) are represented by transcripts that show a complete match with reference sequence (Gff compare class code: =), and thus are defined as known transcripts. However, the majority of DETs identified (1168 out of 1730, 68%) were defined as unknown, as they revealed a partial or none matched with the reference annotation. To go deeper in details, the majority (510 out of 1168, 44%) of unknown or partially unknown DETs, are transcripts which have at least a splice junction shared with the reference sequence (class code j), so they could represent novel isoform of known transcripts. Other well-represented classes of DETs are those constituted by transcripts that are retained in intron(s) and show partial or no intron chain match (class code n; 17%) and by transcripts completely unknown or that map in intragenic

regions (163 out of 1168, 14%) (class code u). The other DETs identified belong to minor represented class code groups, i.e. transcripts that contain the reference sequence (class code k, 8%) DETs that show exonic overlap with the reference but in the opposite strand or a full intron(s) chain match (8%, both class code x and m).

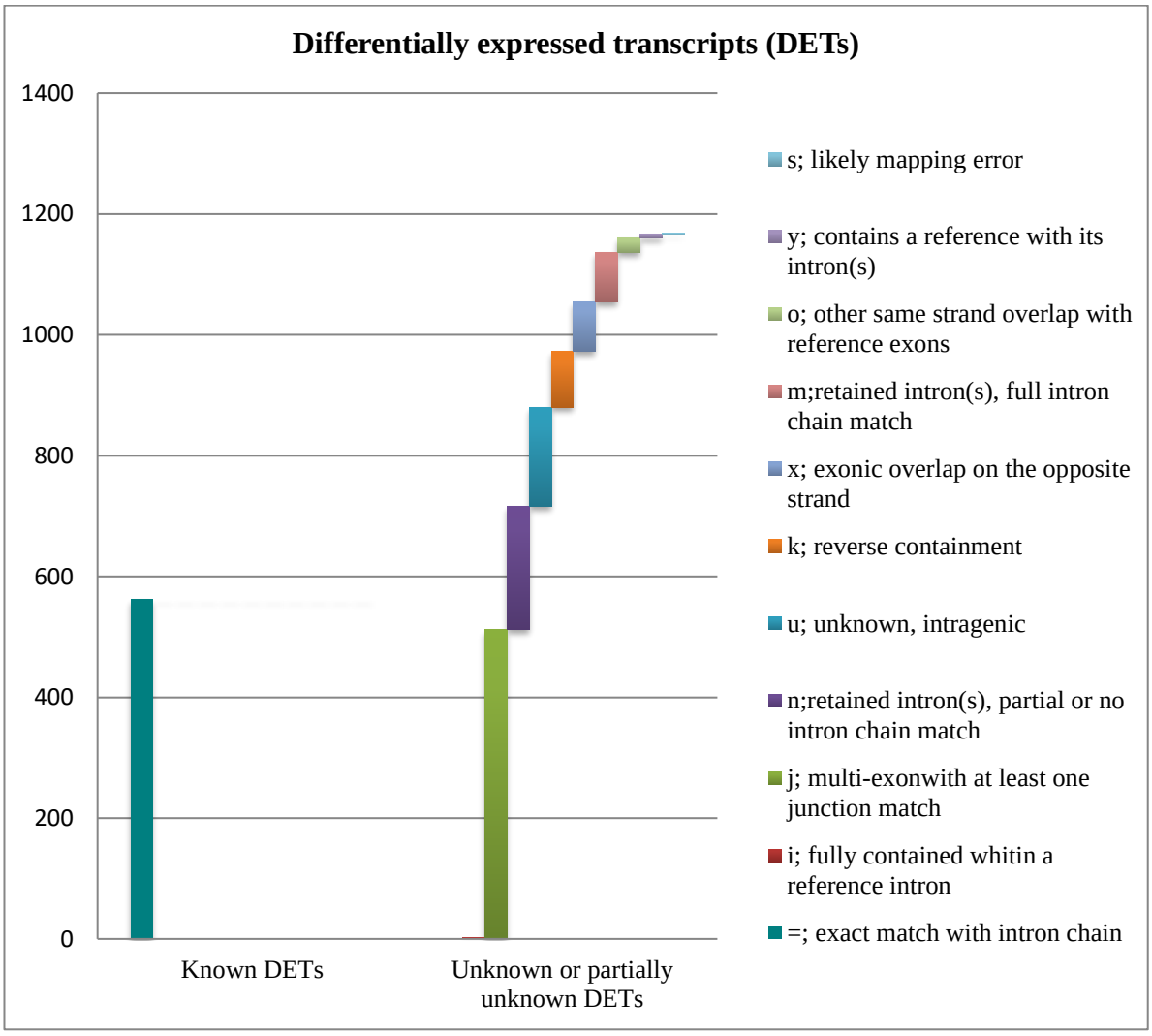


Figure 4.9 DETs' class code: DETs' percentage classified on the basis of Cuff compare's class codes (Ref. Figure 4.8)

In order to better categorize the 562 known DETs identified according to their functional annotations, a pathway analysis was performed. The criteria selected for the identification of pathways were:

- FDR< 0.01
- pathway size range: 5-400 entities

Results provided by Reactome open source tool (Material and Methods section, paragraph 3.5.3), and filtered on the basis of the arbitrary criteria idicated above, the following six biological pathway resulted enriched in the DETs previously identified (3,09E-06 < FDR < 0,0039): endosoma/vacuolar, ER-Phagosome, Antigen Processing-cross presentation, antigen presentation, Insuling-like growth factors binding proteins and collagen chai trimerization pathway (table 4.9). Considering the reference macro area for each pathway identified, three major biological processes emerged to be enriched in DETs: i) the immune system ii) the metabolism of RNA and iii) the extracellular matrix organization. As resported in table 3.10, the immune system macro area includes four out of six pathways identified.

Macro Area	Pathway name	Entities found	Entities total	pValue	FDR
Immune system	Endosomal/Vacuolar pathway	21	82	2,44E-09	3,09E-06
	ER-Phagosome pathway	25	165	1,663E-06	0,0005
	Antigen processing-Cross presentation	25	187	1,379E-05	0,0035
	Antigen Presentation: Folding, assembly and peptide loading of class I MHC	21	102	9,5E-08	6E-05
Metabolism of RNA	Insulin-like Growth Factor-2 mRNA Binding Proteins (IGF2BPs/IMPs/VICKZs) bind RNA	8	13	4,3E-07	0,00018
Extracellular matrix organization	Collagen chain trimerization	11	44	1,9E-05	0,00397

Table 4.9. Pathway analysis on the basis of known DETs. For each pathway identified by Reactome is reported: the macro area, the number of DETs that belong to the pathway (Entities found), the total number of genes that belonges to the pathway (Entities total), p-value and False Discoveri Rate (FDR) relted.

4.3.2 RNA targeted sequencing validation

To confirm the robustness of total RNA sequencing results, DETs identified needed to be validated with another technical approach, both in the same cohort of patients ("technical validation") as well as in an independent cohort with comparable clinical and pathological features ("clinical validation"). Considering the impossibility to exploit qRT-PCR approach to confirm all DETs identified (see Appendix Table 6.1), both for the high number of transcripts and for the impossibility to test the unknown transcripts, targeted RNA sequencing focused on the 1730 DETs was exploited to the aim of this part of the work. In fact, this kind of approach allows not only to test the 562 known transcripts identified, but also to reconstruct the sequence of other DETs that do not show a complete overlap with reference sequence.

Results deriving from the "technical validation" revealed that 286 out 1370 (21%) of DETs tested were as differentially expressed (FDR< 0.05) between Pt-r vs Pt-s patients, on contrary 1084 (79%) were not (figure 4.10 A). Of these 35% (100 out of 286) were known transcripts and 65% (186 out of 286) completely or partially unknown.

Regarding the targeted RNA sequencing validation performed on cohort BI ("clinical validation"), data obtained show that only 82 out of 1370 (6%) of transcripts previously identified are able to discriminate with statistical significance (FDR < 0.05) Pt-r from Pt-s patients in this independent cohort (figure 4.10 B). In particular, 25% (21 out of 82) are known and the remaining 75% (61 out of 82) totally/partially unknown transcripts.

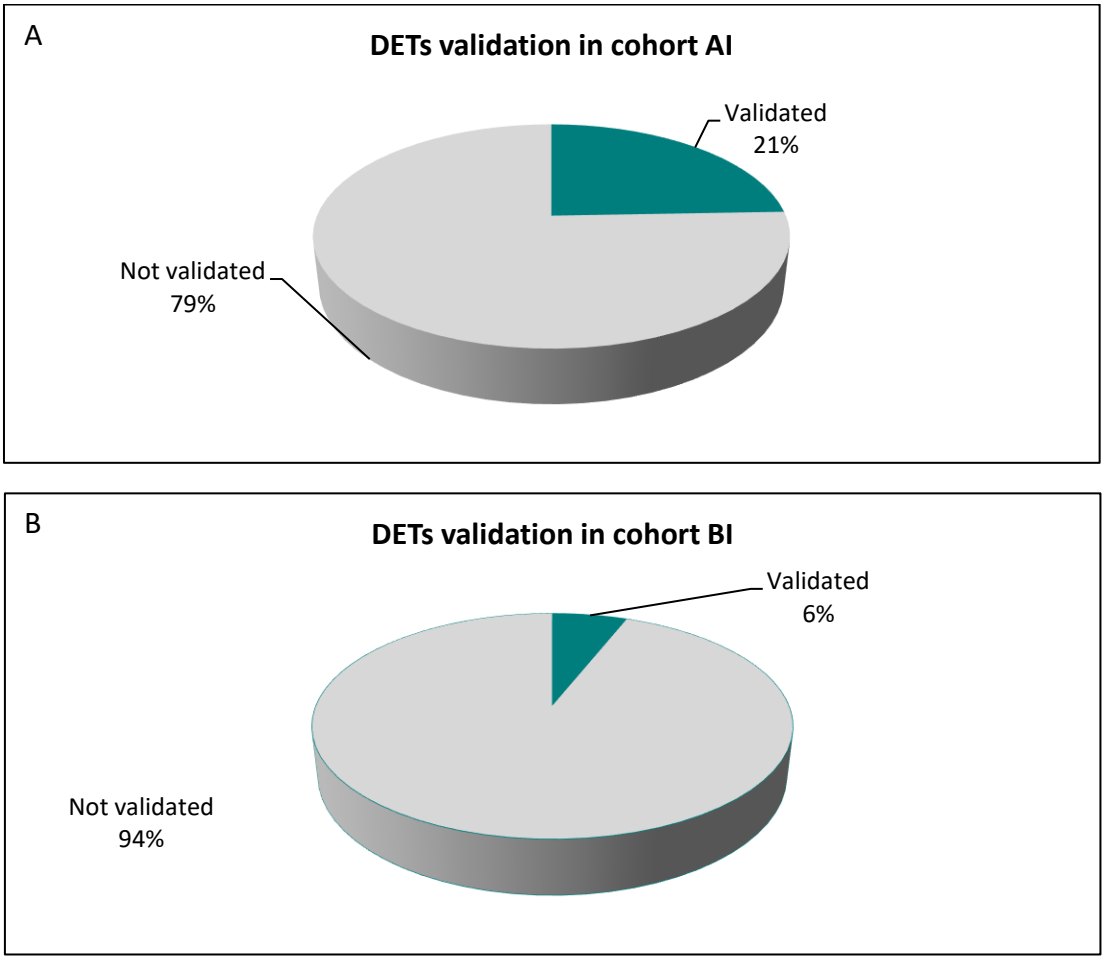


Figure 4.10 Validation of 1730 DETs. Piecharts reporting the percentage of DETs confirmed (dark blue and green color) or not (light blue and green color) as differentially expressed in Pt-r versus Pt-s patients respectively in cohort AI (panel A) and in cohort BI (panel B).

Overlapping results deriving from both "technical" and "clinical" validation, only 25 transcripts out of 1370 (2%) were confirmed in both cohorts of patients with concordant logFC (FDR < 0.05) as significantly differentially expressed between Pt-s and Pt-r patients (Table 4.10).

ref_gene_id	gene_name	class_code	Total RNAseq		Technical Validation Cohort AI		Clinical Validation Cohort BI	
			logFC	FDR	logFC	FDR	logFC	FDR
ENSG00000116299	KIAA1324	j	-2,92	7,93E-05	-2,07	9,95E-06	-1,94	0,007971995
ENSG00000116299	KIAA1324	=	-2,80	0,000213056	-3,14	1,25E-10	-1,85	0,024142495
ENSG00000116690	PRG4	=	2,23	0,008806668	2,37	4,41E-07	1,66	0,021919721
ENSG00000143631	FLG	=	2,84	0,000126688	4,39	3,88E-19	5,49	7,69E-18
ENSG00000276850	CH17-360D5.2	=	2,38	0,002954186	2,43	1,95E-07	2,63	5,84E-05
ENSG00000241170	RNF169	k	-2,49	0,001404577	-1,95	3,44E-05	-1,67	0,026565787
ENSG00000167244	IGF2	=	2,56	0,000830085	1,54	0,001580185	3,55	4,47E-09
ENSG00000166405	RIC3	n	2,26	0,005059663	0,87	0,141887003	6,40	2,75E-22
ENSG00000166405	RIC3	n	2,74	0,000253273	3,93	4,53E-16	1,43	0,064220506
ENSG00000165379	LRFN5	j	2,42	0,005618752	1,59	0,001163736	2,14	0,002630749
ENSG00000279445	OR4N3P	k	4,71	8,99E-11	5,08	7,93E-24	2,65	1,85E-05
ENSG00000166426	CRABP1	=	-2,86	0,000108584	-3,37	1,50E-12	-3,89	3,15E-09
ENSG00000259424	RP11-35O15.1	o	3,33	4,56E-06	2,70	7,42E-09	1,90	0,009577545
NA	MUC16	x	3,19	5,31E-05	1,22	0,017601886	1,61	0,026565787
NA	NA	u	2,20	0,009617488	1,39	0,008134576	3,47	1,32E-05
ENSG00000261824	LINC00662	=	2,26	0,006614098	1,19	0,020372608	2,49	7,00E-05
ENSG00000233296	AC092159.2	j	-4,59	2,26E-10	-2,69	8,10E-09	-3,04	3,64E-06
NA	NA	u	3,03	0,000118988	3,40	1,68E-12	1,82	0,04260844
ENSG00000125872	LRRN4	=	2,68	0,000377228	2,59	2,82E-08	3,79	4,59E-10
ENSG00000239311	RP11-615J4.3	j	8,69	1,07E-25	5,42	4,21E-26	1,87	0,006609068
ENSG00000206503	HLA-A	=	-2,26	0,004906301	-1,32	0,008353072	-1,51	0,053752143
ENSG00000164692	COL1A2	j	2,33	0,003170749	1,20	0,01892756	2,62	3,31E-05
ENSG00000091656	ZFHX4	m	2,41	0,002172487	2,19	2,83E-06	2,46	0,000110018
ENSG00000255530	ZNF252P	k	-2,95	0,000113404	-1,44	0,003560485	-5,74	1,34E-16
NA	NA	u	3,78	1,23E-07	0,96	0,084167842	3,44	1,34E-08

Table 4.10 DETs confirmed with targeted RNAseq in Cohort AI and BI. For each DET confirmed is reported the Ensemble reference gene ID (ref_gene_id), gene name, class cose (gffCompare), the relative logFC and FDR derived from total RNAseq analysis and targeted RNAsequencing analysis in Cohor AI (technical validation) and Cohort BI (clinical validation).

Column chart reported in figure 4.11 shows that the majority of DETs confirmed (36%, 9 out of 25) are represented by known transcripts, which show a complete match with intron chain, whereas three are completely unknown. The remaining 13 DETs are represented by potential new isoforms of known transcripts (five out of 25), by transcripts that contain the reference sequence (three out of 25) or by transcripts that show a partially or total overlap with intronic/exonic regions (five out of 25).

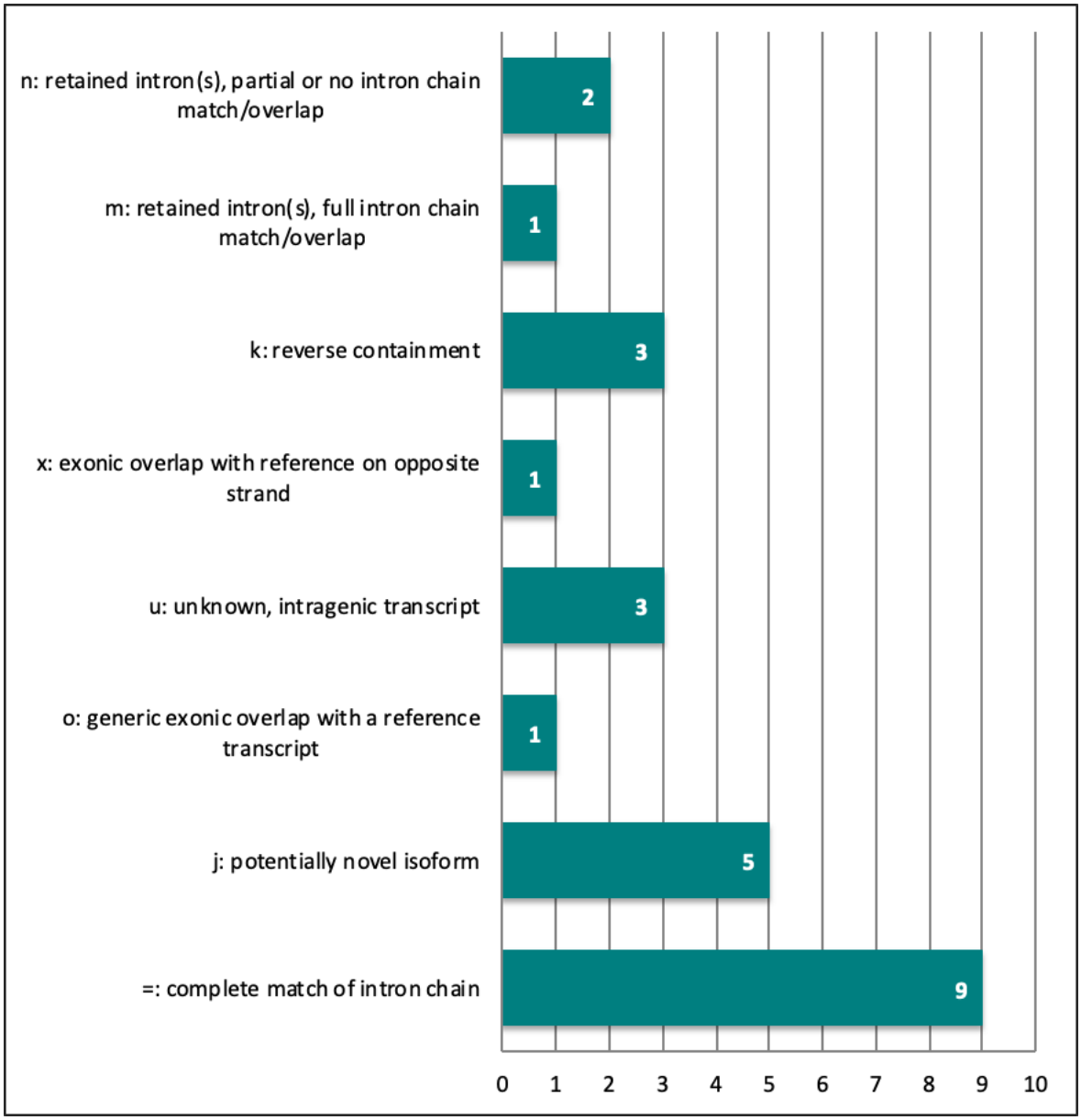


Figure 4.11 Validated DETs' class codes. The 25 validated DETs were subdivided on the basis of their Gff compare class code. For each class the number of its associated DETs is reported

The nine known transcripts identified and confirmed as DETs between Pt-r and Pt-s patients are constituted by eight transcripts derived from protein coding genes and one transcripts (LINC00662) that is a long intragenic non-protein coding RNA (lincRNA). Of these eight DETs, three are down-regulated and five up-regulated in Pt-r versus Pt-s HGS-EOC cases (table 4.11).

DETs with complete match with intron chain (class code: =)			
Chromosome	**Corresponding Gene Name**	**Type of gene**	**Expression in Pt-r vs Pt-s**
1	*KIAA1324*	Protein coding	Down-regulated
1	*PRG4*	Protein coding	Up-regulated
1	*FLG*	Protein coding	Up-regulated
10	*CH17-360D5.2*	Protein coding	Up-regulated
11	*IGF2*	Protein coding	Up-regulated
15	*CRABP1*	Protein coding	Down-regulated
19	*LINC00662*	lincRNA	Up-regulated
20	*LRRN4*	Protein coding	Up-regulated
6	*HLA-A*	Protein coding	Down-regulated

Table 4.11 List of validated DETs with complete match with intron chain. For each of the eight DETs validated is reported the relative chromosome, the name of corresponding gene, the type of gene and the differential expression in Pt-r versus Pt-s HGS-EOC patients.

4.3.3 Conclusion

In this third part of the Results section, the differences in trascriptomic profile between Pt-r and Pt-s HGS-EOC patients were investigated by RNA sequencing approaches.

As overall technical comment, our results demonstrated that a double-step RNA sequencing approach, consisting in total RNA sequencing followed by a targeted RNA sequencing both in the same and in an independent cohort of patients, is necessary to obtain robust and reproducible data. In fact, considering the targeted RNA sequencing validation, the percentage of DETs not

validated in the same cohort of patients (AI) is 79% and further increased considering the results obtained from the independent cohort (BI), where the percentage of DETs confirmed is only 6%. Moreover, overlapping data derived from cohort AI and BI, only 25 out of 1370 transcripts (2%) were confirmed as differentially expressed between Pt-r and Pt-s cases.

Among those only nine out 25 (36%) are represented by transcripts derived from coding genes, confirming the limits of array-based approaches in investigating the fine transcriptional regulatory mechanisms underlying Pt resistance. Results presented in fact, clearly demonstrated that 64% of DETs identified are partial or completely unknown transcripts, that could not have been identify with the array technology.

A functional investigation of elements composing this transcriptional signature, and the integration of these data with other independent Pt-predictive biomarkers, could allow in future to efficiently stratify patients at time of diagnosis and to better understand biological mechanisms that characterized the complex scenario related to Pt resistance in HGS-EOC.

5. Discussion and Conclusion

The identification of molecular signatures that recapitulate the main biological discriminants between Pt-s and Pt-r cases with potential clinical utility has been the main goal of many studies conducted over the last decades. The results reported in this work of thesis would represent one step forward in the deciphering the complex molecular puzzle of Pt resistance in HGS-EOC.
Considering the three aspects investigated and mentioned above, the discussion is organized in three main sections, at the end of which the future prospectives for this study are illustrated.

5.1 Pathway analysis

In the first part of this research project a three genes-based signature *(PRKG1, SDF2L1,* and *PPP1R12A),* called *SIII* signature, that reflects the biology of primary Pt-r tumors and predicts prognosis in HGS-EOC patients, was identified.

The comparative transcription analysis of Pt-s and Pt-r cancer patients, in keeping with some previous reports (133), allowed us to identify a restricted number of mRNAs and miRNAs the expression of which discriminates between sensitive and resistant cases. The confidence of these findings was strengthened by the validation of 18 over 23 (78%) selected genes in a different cohort of cases from an independent tissue biorepository as well as in publicly databases, included TCGA.

The core of our pathways wired 131 mRNAs and five miRNAs, with the *PRKG1* gene playing a role as a hub of the entire network. In our network miRNAs do not seem to play a central role. It is difficult to explain the reasons for this result since it could be due to many factors from technical sensitivity to biological effects.

miRNAs are modulatory molecules for which small expression changes might have big post-transcriptional impact on many target genes, acting as mediators of different pathways. In this perspective we can speculate that their small modulations combined with the large biological variability of the tumour at least at the time of diagnosis could contribute to mask the miRNA effects in primary resistance. It is worth noting that some non-coding elements of the network

have been previously identified as prognostic in HGS-EOC. For example, *miR-193-5p* -highly expressed in Pt-r patients- is an element of the prognostic MIROvaR signature (95).

Since primary resistance against Pt-based chemotherapy is associated with poor prognosis, we investigated the prognostic role of the *SII* signature in one of largest available databases of EOC expression profiles (the Curated Ovarian Database - cohort C).

It is important to highlight that these retrospective databases were not originally intended to be used for studies of chemo-resistance as the primary purpose. We realized that the use of different gene expression technologies as well as metrics for optimal treatment response might introduce potential bias in the data and in the analyses. To minimize this bias and to reduce the impact of intra-patient tumor heterogeneity, we used very stringent statistical criteria to limit the risk of false positive predictive genes. This strategy identified three genes, *PRKG1, SDF2L1* and *PPP1R12A, alias SIII* signature -as prognostic biomarkers in HGS-EOC. Moreover, the combination of the expression profiles of *SIII* signature significantly predicts patient prognosis with a synergic effect.

PRKG1 (chr: 10q11.23) is a cyclic GMP (cGMP)–dependent protein kinase (PKG), poorly characterized in terms of biological function, in particular in HGS-EOC. Recently the cGMP/PKG signaling pathway was found to play an important role as an antiapoptotic mechanism in ovarian cancer cell lines, by promoting cell survival, through interaction with the *src* gene. Moreover, the NO/cGMP/PKG signaling pathway has been reported to protect human ovarian cancer cells against both spontaneous and cisplatin-induced apoptosis (134).

The *SDF2L1* gene (chr: 22q11.21) encodes for a member of the stromal cell derived factors family (SDF) secreted by stromal cells, including fibroblasts. Our results are consistent with the evidence that low levels of *SDF2L1* are associated to poor prognosis, relapse and metastasis in breast (135), colon (136) and ovarian cancer (137).

PPP1R12A (protein phosphatase 1 regulatory subunit 12A) (chr: 12q21.2-q21.31) belongs to the myosin phosphatase targeting protein (MYPT) family. It is also known as myosin phosphatase target subunit 1 (MYPT1). In cancer cells, PPP1R12A plays a critical role in major regulatory pathways such those pertaining to Wnt/β-catenin signaling (138) and PI3K/AKT (139). In has also been demonstrated that PPP1R12A is a key regulator of the Hippo pathway, and that triggered by external stimuli like EGF or TGF1 it modulates the nuclear localization and stability of YAP/TAZ, thus controlling cell growth, proliferation and EMT (140). Despite this important mechanistic role, few studies address the potential prognostic role of *PPP1R12A.* Genome instability in the P*PP1R12A* locus gene has been found as independent predictor of recurrence and overall survival in colorectal cancer patients receiving oxaliplatin-based adjuvant chemotherapy (141). Data as to its predictive/prognostic value in HGS-EOC has hitherto not been published.

It is plausible to hypothesize that *PRKG1, SDF2L1, PPP1R12A* genes are part of a larger prognostic signature, the elements of which have not been identified here due to the stringent statistical thresholds applied and to the use of non-uninform technical platforms for data generation. For example, we have previously identified *miR-181a-5p* as positive regulator of TGFβ and EMT in HGS-EOC, and its expression levels are independent prognostic factors of survival [32]. In our network we did not identify *miR-181a-5p* but rather STAT1 and STAT3, the former a modulator and the latter a target of *miR-181a-5p*. PPP1R12A has been reported to be activated by TGFβ a regulator of EMT, through *miR-181a-5p* (94), confirming the importance of the EMT in the control of both, resistance against platinum and response to therapy of HGS-EOC.

5.2 Genomic instability analysis

The second result of this work of thesis derives from the correlation analysis between defects in Homologous Recombination pathway and sensitivity to DNA-damaging agents, like Pt compounds, which has become over the last years a key clinical relevant paradigm for HGS-EOC treatment (142) (124). Initially the deficiency in HR pathway was defined only on the basis of the detection of germline or somatic mutations in *BRCA1* and *BRCA2* genes. For example previous studies have demonstrated a higher improvement in sensitivity to Pt-based therapy and also PARPi of ovarian cancer patients carrying *BRCA1/2* mutations relative to *BRCA1/2* wild-type tumours (142) (143).

However, alterations that affect *BRCA* genes explain only the 50% of HR defects, indicating that alterations in other HR-related genes are responsible for deficiency in HR pathway (144)

Whole-genome sequencing analysis following the trinucleotidic model proposed by Alexandrow and colleagues (145) has identified a specific mutational signature - "signature 3"- that was strongly associated to bi-allelic inactivation of *BRCA1* and *BRCA2* alterations and that was also able to identify HRD phenotype in a further 16% of cases not characterized by canonical HRD related defects (146). This approach would be useful to stratify HRD positive and negative patients, however its application in clinic seems to be limited both for the complex bioinformatic analysis required to generate the trinucleotidic signatures and for the costs of the experiments per patient.

Thus, to develop a new approach applicable in clinical setting, the NGS assays to determine the HR status, were becoming focused on evaluation of the presence of large genomic effects induced by the genomic instability and on the mutational analysis of a minor number of target region.

The two commercial tools to identify a genomic scar (see Inroduction, Section 1.4.2.2) are: i) "myChoice®" (Myriad) that tests for the presence of the three main structural genomic

alterations (loss of heterozygosis, telomeric allelic imbalances, large scale transitions) and ii) the "FoundationFocus CDx $_{\text{BRCA LOH}}$" (FoundationOne®) which detects the presence of mutations in *BRCA1/2* genes and the percentage of genome affected by LOH.

Using these assays, HRD status was investigated in many ovarian cancer clinical trials to select patients who could benefit from PARPi treatment (47) (147) (148), or in triple negative breast cancer trials to predict response to DNA-damaging agents or to evaluate the addition of carboplatin to standard chemotherapy (124) (149).

To create a patent-free independent academic test to evaluate the status of HR pathway in the clinical setting, we started to develop an in-house HRD assay. Results reported in this thesis work represented an initial step for the setting of the assay that will be implemented and compared with the clinical commercial gold-standard for the evaluation of HR deficiency (MyChoice®, Myriad).

In line with MyChoice® assay, we evaluated the sum of the three main genomic scars and, as described in Results section, we set a threshold of 52 to distinguish HRD positive (>52) from HRD negative (<52) patients. HRD analysis reported in this thesis work is preliminary and at the moment it suffers of three main limitations:

i) As the limited number of *BRCA* germline mutated patients (n=10), that would represent our groundtruth set, the construction of ROC curves to validate the HRD score threshold decided was not feasible in this thesis work. ROC curves would allow to test different thresholds in order to establish the one associated with best values of specificity and sensitivity. To achieve this goal, it is planned to recruit at least 50 new *BRCA* germline mutated patients to have a robust groundtruth set (true positive cases) of comparison.

ii) At the moment, the HRD in-house analysis has considered only the three genomic scars investigated by "MyChoice®". To try to improve the sensitivity of the test proposed it is planned to integrate the evaluation of LOH, TAI and LST with

mutational information regarding HR-related genes and genes involved in Pt and PARPi resistance, such as ABCB1.

iii) The academic test proposed has been developed and tested using the same cohorts of patients, but it has already planned to verify its predictive value on a larger and totally independent cohort.

The integration approach that will be developed in-house for the identification of HGS-EOC patients carrying defective in homologous recombination pathway would allow us to better identify patients who could benefit not only from PARPi treatment, but to give a contribution in the developing of a genetic/transcriptomic signature able to stratify patients on the basis of Pt sensitivity at time of diagnosis.

5.3 Transcriptomic analysis

As previously stated, a great number of transcriptional signatures associated to Pt-based therapy have been developed using array technology, which allowed to focus only on the known part of the transcriptome.

However, despite a huge amount of data has been generated, the mechanisms underlying Pt resistance have not been fully understood yet, suggesting that for a more detailed comprehension of this complex mechanism requires a fine analysis across the entire human transcriptome including known and partially or totally unknown transcripts.

From technological point of view, the advent of NGS technology, which allows the absolute quantification of transcripts abundance, the identification of novel transcripts' isoforms, or transcripts species has become over the last years the best high-throughput approach for a comprehensive fine mapping of the entire transcriptome.

Up to now, the majority of studies published in literature have used the RNA sequencing technology to dissect the mechanism of Pt resistance in HGS-EOC – i.e. in cellular models

(150) (151), in patients-derived xenografts (PDXs) (152) and in HGS-EOC patients' biopsies (153) (154)- focusing only on the absolute quantification of known transcripts, but no attention was directed to the identification and the investigation of the role of novel transcripts.

In this work of thesis, we used a high-coverage total RNA sequencing approach to achieve absolute quantification of known transcripts as well as to discover and to evaluate the expression levels of transcripts that have never been described before.

To reinforce data obtained with total RNA sequencing, thus avoiding the problem of false positive data, we exploited a targeted approach on one hand to confirm the differentially expression of known DETs identified and on the other hand to validate both the structure and the expression of reconstructed transcripts. This double-step RNA sequencing approach has never been described in literature before and would make more reliable and robust the data that were obtained. In fact, it is important to note that the validation approach applied did not confirm the 79% -in the same cohort- and the 94% -in an independent cohort of patients- DETs identified in the first step of RNA sequencing analysis.

In our results, considering the validated 9 known DETs identified, we found that three of these resulted down-regulated in Pt-r compared to Pt-s patients. Among those KIAA1324 – also known as EIG121-, that encodes for a transmembrane protein whose expression is induced by estrogen. Although its biological function is still poorly understood, the expression levels of KIAA1324 is reported to be correlated with good prognosis in endometrial (155), pancreatic (156) and ovarian cancer patients. Although we did not perform a survival analysis, in line with our results which indicate an overexpression of KIAA1324 in Pt-s patients, the analysis of RNA sequencing data derived from 373 ovarian cancer patients –TCGA dataset-, suggests that high expression of KIAA1324 is significantly associated with favorable prognosis (cut-off: 0.26, p-score 0.0002) (157). Another DET that is down-regulated in Pt-r cases is the one that encodes for CRABP1 protein, which is a specific binding protein for a vitamin A family member and is thought to play an important role in retinoic acid-mediated differentiation and proliferation

processes. CRABP1 has been associated with poor prognosis in triple-negative breast cancer and it is demonstrated to inhibit the cellular growth arrest induced by retinoic-acids (158). In line with our results, the survival analysis performed by Miyake and colleagues (159) demonstrated that reduced CRABP1 expression is significantly associated with poor OS and PFS in both serous and clear cell ovarian carcinoma. Since CRABP1 is selectively expressed by mesenchymal tissues at the junction of the epithelium and the mesenchyme (160), its reduction in ovarian surface/Fallopian tube cells could lead to a de-differentiation of cells themselves and a consequent EMT process activation, that it is known to be related to Pt resistance (161) (92). The last known DETs that from our results emerged as down-regulated in Pt-r patients is HLA-A, which belongs to the HLA class I complex. Although we did not perform survival analysis in this part of the work, it was documented that a down-regulation of HLA class I factors is one of the mechanisms involved in tumour escape from the immunological surveillance (162) and that this down-regulation is associated with poor outcome in ovarian cancer patients (163).

Among the six DETs that resulted up-regulated in Pt-r versus Pt-s HGS-EOC, to our knowledge no confirmed evidences regarding a possible role in Pt resistance or in HGS-EOC biology is reported for the proteoglycan 4 (PRG4), FLG -a filament-associated protein that aggregates keratin filament in mammalian epidermis- for ANXA8 (CH17-360D5.2) which encoded for a protein with anticoagulant activity whose overexpression is reported to be associated with poor prognosis in early stages of pancreatic cancer (164), and for LRRN4 –leucine rich repeat neuronal 4-. Another known transcript that from our analysis resulted up-regulated in Pt-r HGS-EOC patients is IGF2. IGF2 –insulin like growth factor 2- is a member of the insulin family of polypeptide growth factors, which through the interaction with IGF1R –insulin growth factor receptor 1-, lead to the activation of oxygen-independent pathway promoting the transcription activity of HIF-1 (Hypoxia Inducible Factor 1). HIF-1 activated the transcription of many genes that encode for proteins that are involved in biological processes like angiogenesis, migration, cell survival and invasion, resistance to therapy in many cancer types including triple negative

breast cancer (165), esophageal cancer (166), osteosarcoma (167) (168). In ovarian cancer it has been reported significant association between high levels of IGF2 mRNA and shorter OS and PFS and a possible role of IGF2 in resistance to taxol treatment (169). It is important to note that the activation of HIF-1 promotes the transcription of several genes –including Twist-1, Snail and ZEB1/2, that are related to EMT activation (170) and thus could be involved in Pt resistance (Introduction section, paragraph 1.4.3). Although its role is not clearly elucidated yet, the long intragenic non-coding RNA 00662 (linc00662) that we found overexpressed in Pt-r patients compared to Pt-s ones, has been reported to be associated with cell growth in gastric cancer cells (171) and with OS in lung cancer patients (172). To our knowledge no information regarding its possible involvement in therapy resistance is reported in literature.

To sum up, it is evident that at the moment no data are available in literature on the role and the involvement of these 9-known transcripts in Pt resistance phenomenon, thus the mechanisms by which the expression of each of these 9 transcripts is associated with intrinsic resistance to Pt remains to be elucidated in future studies.

5.4 Future prospectives

In this work of thesis, the complex and heterogeneous phenomenon of Pt resistance in HGS-EOC has been investigated through different genomic and transcriptomic approaches, providing initial insights into the changes in gene expression and alterations in HR repair system associated with Pt resistance. In perspective these findings may contribute to improve early identification of patients who would benefit from Pt-based therapy from those who would not, thus avoiding ineffective toxic treatment.

The ultimate aim of this work is to identify a reliable predictive signature applicable in the clinical setting. Based on the results obtained in the present thesis, it is planned to conduct further research aimed to:

1. To integrate the two transcriptional signatures presented in paragraphs 4.2 and 4.4 of the Results. At present these signatures cannot be integrated as they were generated by different methods. They will be tested using the same technique (targeted RNA sequencing) on a cohort of FFPE HGS-EOC tumour samples.
2. To improve the in-house HRD assay (see paragraphs 4.3 of the Results) by including the mutational analysis of genes which are reported to be associated to Pt-resistance or involved in HR pathway, using the same FFPE HGS-EOC tumour samples as in point 1.
3. To compare the integrated transcriptional signature obtained in point 1 with the implemented HRD assay developed in point 2, to identify the best approach to discriminate at time of diagnosis Pt-sensitive from Pt-resistant patients, allowing to apply the best therapeutic option.

6. Appendix

Table 6.1: **Differentially expressed transcripts derived from total RNA sequencing analysis.**

For each DETs identified are reported the chromosome location, log Fold Change (logFC), False Discovery Rate (FDR), the Gff compare class code and, where possible, the nearest reference gene.

chr	logFC	FDR	class_code	gene_name
14	-8,725	4,71E-20	j	DICER1-AS1
14	-8,439	1,59E-20	k	IGHG1
1	-7,985	2,28E-17	=	NRD1
16	-7,825	3,40E-17	n	PKD1P1
4	-7,741	3,11E-16	j	TAPT1-AS1
12	-7,731	6,90E-16	j	TMTC2
17	-7,475	9,57E-16	j	SPAG9
11	-7,466	2,77E-19	n	PPP6R3
13	-7,403	6,40E-20	j	TMCO3
2	-7,317	1,45E-15	m	LANCL1
18	-7,281	3,84E-15	=	LINC00908
11	-7,252	5,24E-15	=	EPS8L2
10	-7,189	1,58E-14	j	SORCS1
15	-7,073	2,31E-14	j	POLG
12	-6,947	1,74E-10	x	A2ML1-AS2
16	-6,932	1,55E-13	n	CDK10
22	-6,871	1,32E-10	j	GCAT
10	-6,863	6,10E-19	m	NRG3-AS1
3	-6,793	1,57E-12	n	ADAMTS9
9	-6,673	9,63E-10	j	ZNF782
5	-6,438	8,99E-11	j	ANKHD1-EIF4EBP3
17	-6,386	1,21E-11	j	SGSH
14	-6,327	1,11E-16	j	RP11-666E17.1
3	-6,312	4,66E-14	j	FNDC3B
16	-6,182	4,93E-11	n	MAPK8IP3
9	-6,173	4,66E-07	k	ANKRD18B
1	-6,172	2,75E-05	=	FCGR3B
12	-6,156	1,78E-10	k	CPM
7	-6,065	8,33E-06	j	POLR2J2
2	-6,060	1,85E-10	m	NR4A2
12	-6,046	3,17E-10	=	NELL2
12	-6,030	4,62E-10	j	ATXN2
X	-5,979	5,95E-05	j	RAP2C-AS1
16	-5,957	4,24E-10	=	CARHSP1
11	-5,854	9,92E-10	n	PDZD3
14	-5,772	3,72E-09	j	SLC25A29

15	-5,767	1,56E-09	=	SMAD3
16	-5,728	1,93E-09	j	RP11-505K9.4
5	-5,721	1,80E-07	=	FAM153B
4	-5,701	1,93E-08	=	RAPGEF2
12	-5,691	4,04E-09	=	PFKM
17	-5,667	7,83E-06	n	MYO15B
1	-5,641	5,52E-09	n	RP11-504P24.3
14	-5,633	1,36E-08	m	RP11-111A21.1
8	-5,603	0,000422788	m	STAR
11	-5,602	1,78E-09	m	LRP5
15	-5,545	1,33E-08	j	GOLGA8B
20	-5,542	1,27E-07	=	ABHD12
1	-5,521	1,90E-08	j	FGR
12	-5,492	2,52E-08	=	PHLDA1
7	-5,481	1,01E-06	k	ARHGEF34P
4	-5,476	1,38E-08	=	BST1
21	-5,475	5,54E-08	j	MCM3AP
15	-5,430	5,16E-08	j	RP11-661P17.1
12	-5,405	1,08E-07	o	RP11-446N19.1
19	-5,361	2,59E-08	=	FOSB
6	-5,353	3,18E-07	=	RP11-527F13.1
16	-5,342	4,52E-08	j	MYLK3
17	-5,315	5,03E-08	j	LIMD2
19	-5,312	1,23E-07	=	NFIX
14	-5,303	5,67E-08	j	NUMB
17	-5,299	1,35E-05	=	ACE
X	-5,283	0,00069139	j	NHSL2
1	-5,278	1,08E-07	=	NFIA
14	-5,271	3,84E-10	m	TRAV8-3
6	-5,264	4,59E-05	=	TNFAIP3
6	-5,251	2,26E-12	x	RP1-80N2.2
4	-5,238	1,21E-10	n	SH3BP2
6	-5,152	7,07E-06	=	MCM9
12	-5,146	3,09E-07	j	LMNTD1
17	-5,137	3,51E-10	k	MYO18A
X	-5,130	0,000171588	j	DMD
3	-5,113	4,57E-07	=	ZBBX
7	-5,112	3,45E-07	n	INSIG1
11	-5,074	1,92E-10	n	TTC17
6	-5,071	7,90E-05	k	RP3-329A5.8
3	-5,070	1,46E-06	j	ATXN7
2	-5,070	4,43E-08	=	LINC00570
9	-5,060	3,76E-06	j	RFX3-AS1
2	-5,045	6,21E-07	=	LIMS1
5	-5,021	3,12E-08	j	ZNF354B
6	-5,009	2,39E-10	j	PHF3
3	-5,001	3,19E-06	=	EIF4G1

2	-4,985	5,80E-11	j	ERBB4
1	-4,950	1,21E-10	n	ABL2
2	-4,947	8,99E-07	n	PPP1R21
17	-4,930	1,40E-06	j	SKA2
4	-4,920	1,73E-06	x	RNF150
14	-4,907	4,20E-07	=	LRRC9
22	-4,905	1,01E-06	j	FAM118A
11	-4,888	7,80E-07	j	CEP164
6	-4,884	5,11E-07	j	EZR
18	-4,872	1,46E-11	=	ONECUT2
3	-4,841	1,04E-05	=	TNIK
6	-4,838	9,77E-08	j	PHF3
5	-4,816	4,56E-06	j	IQGAP2
8	-4,808	0,005676189	=	ANK1
10	-4,801	1,92E-07	n	FAM208B
15	-4,773	2,88E-06	j	EFTUD1P1
X	-4,734	4,48E-05	j	RP11-761E20.1
17	-4,723	3,44E-06	j	FOXK2
2	-4,719	2,39E-06	m	CATIP
X	-4,705	0,000798941	j	DDX26B
11	-4,705	4,67E-06	j	ANO1
20	-4,670	2,80E-08	u	NA
9	-4,661	0,008344115	=	C9orf3
5	-4,617	0,000220937	=	TCF7
4	-4,616	3,26E-06	=	TMEM175
7	-4,600	6,65E-05	=	RP11-448A19.1
15	-4,600	7,49E-06	m	ARRDC4
2	-4,591	2,26E-10	j	AC092159.2
5	-4,587	1,01E-05	=	TNFAIP8
11	-4,560	4,19E-09	u	NA
3	-4,554	9,00E-06	=	CFAP44
6	-4,536	0,000311146	=	GABBR1
22	-4,513	0,000104043	j	TANGO2
2	-4,490	2,31E-07	=	AC093838.4
17	-4,489	2,60E-08	n	TBX2
22	-4,473	4,95E-09	=	APOL4
9	-4,471	0,000110707	j	RP11-500B12.1
21	-4,469	3,07E-09	=	TFF3
5	-4,463	8,33E-06	j	WWC1
4	-4,459	2,71E-05	u	NA
3	-4,441	4,00E-08	j	MRPS22
3	-4,434	4,39E-05	=	TRAK1
18	-4,428	1,96E-05	j	ZBTB7C
11	-4,423	1,81E-05	=	HMBS
6	-4,421	4,54E-05	j	RNF217
17	-4,409	2,47E-05	j	NOL11
20	-4,404	1,41E-09	j	MIR646HG

7	-4,403	9,51E-05	x	RP11-397J20.1
14	-4,392	1,48E-07	j	SPTLC2
22	-4,390	3,72E-05	j	LARGE
4	-4,387	0,000435009	=	SH3RF1
3	-4,381	2,98E-09	j	NLGN1
17	-4,381	1,86E-09	u	NA
7	-4,376	0,000139554	x	AVL9
5	-4,369	5,34E-05	j	MAST4
13	-4,368	3,24E-05	j	KIAA0226L
2	-4,355	2,01E-05	=	SCTR
19	-4,344	3,27E-05	j	NFIX
1	-4,343	1,80E-07	k	WASH7P
17	-4,332	3,12E-05	j	LIMD2
21	-4,315	0,000289278	j	ITGB2
2	-4,310	6,22E-07	j	ERBB4
19	-4,294	4,09E-05	j	OAZ1
1	-4,292	5,66E-08	j	LINC00970
1	-4,287	1,23E-07	j	GPR157
1	-4,279	2,69E-09	n	SUCO
16	-4,268	4,18E-08	u	NA
18	-4,247	1,66E-07	o	DLGAP1
11	-4,243	3,32E-09	j	FTH1
7	-4,229	0,002223803	j	RSPH10B2
10	-4,227	8,45E-05	j	C10orf107
16	-4,198	6,27E-08	u	NA
6	-4,198	0,000153413	j	PM20D2
2	-4,197	9,36E-07	u	NA
10	-4,196	2,59E-07	j	PFKP
2	-4,165	7,90E-05	j	WIPF1
8	-4,164	0,00014098	j	RBPMS-AS1
19	-4,153	0,000330667	j	FKBP8
21	-4,151	8,04E-09	=	TFF3
12	-4,151	3,46E-08	m	GCN1L1
11	-4,150	9,04E-05	j	CEP164
7	-4,149	0,006279762	j	WDR86-AS1
11	-4,125	5,27E-06	=	CCDC88B
7	-4,109	0,000102508	=	PIP
X	-4,106	3,10E-08	u	NA
16	-4,105	2,94E-08	n	TP53TG3D
5	-4,103	7,61E-05	=	AFAP1L1
7	-4,103	0,000469004	j	DPY19L2P1
1	-4,094	1,70E-07	=	TMEM56
19	-4,087	0,000377228	j	LMNB2
3	-4,077	0,000283181	x	LPP
1	-4,063	1,81E-05	u	NA
11	-4,059	1,51E-08	k	RP11-428C19.5
20	-4,059	0,000542566	=	NFATC2

15	-4,056	2,76E-05	=	ACSBG1
20	-4,044	7,14E-06	n	CTSA
X	-4,002	8,45E-06	u	NA
11	-3,977	0,000278679	=	CAPRIN1
12	-3,970	3,27E-07	u	NA
22	-3,966	0,000774632	j	KIAA1671
16	-3,963	0,000231426	j	ZNF778
5	-3,959	0,005756042	n	HK3
1	-3,951	6,21E-08	=	NBPF14
5	-3,921	0,001172578	j	ADAMTS12
15	-3,915	2,59E-07	k	GOLGA8O
12	-3,911	0,000283702	=	MLXIP
6	-3,910	1,39E-07	=	PKHD1
2	-3,903	2,67E-07	=	RAPGEF4
5	-3,890	0,002568939	=	CCNO
22	-3,887	2,99E-05	j	FAM118A
3	-3,885	9,66E-08	i	RP11-441M10.1
5	-3,881	0,001230094	=	FAM153B
15	-3,863	0,000880388	=	CYP19A1
4	-3,855	0,000520918	=	CNGA1
1	-3,840	1,74E-07	n	HIVEP3
10	-3,838	0,000677138	x	ARID5B
10	-3,821	0,000579324	j	HK1
X	-3,801	0,001236153	n	CXorf22
11	-3,789	1,25E-05	j	C11orf70
12	-3,777	1,35E-05	k	WASH7P
6	-3,770	1,40E-07	=	CRISP3
19	-3,765	0,001668413	=	NR1H2
21	-3,760	5,53E-06	n	MCM3AP
2	-3,757	6,25E-05	j	ERBB4
5	-3,744	0,000911355	j	GRAMD3
22	-3,740	0,00011318	j	MICAL3
3	-3,737	0,001392882	j	FOXP1
3	-3,726	1,35E-05	n	ARIH2
21	-3,718	3,16E-06	=	PRDM15
2	-3,713	3,02E-07	j	ZDBF2
X	-3,699	1,71E-05	u	NA
2	-3,688	0,000937582	j	AC109826.1
16	-3,686	0,003222421	n	SMPD3
X	-3,679	3,58E-06	x	RP11-305F18.1
1	-3,660	0,002384112	=	EPHA2
12	-3,652	0,002799082	j	RP11-256L6.3
8	-3,640	0,004504678	k	RP11-1081K18.1
12	-3,633	0,000927131	n	ENO2
17	-3,631	0,001393341	j	CDC6
20	-3,628	6,73E-07	m	RP5-1016A21.1
2	-3,616	0,001810263	j	DNAH7

1	-3,614	0,001748262	k	RP11-275I14.4
19	-3,614	3,51E-05	=	CEACAM21
20	-3,602	5,35E-07	j	MIR646HG
12	-3,582	0,002380527	n	GOLGA2P5
14	-3,574	1,57E-06	n	CTD-2566J3.1
20	-3,568	0,006189716	u	NA
8	-3,568	6,90E-06	=	AP3M2
18	-3,552	1,01E-06	=	MC4R
20	-3,552	8,11E-07	=	SYCP2
1	-3,543	1,52E-06	=	YY1AP1
11	-3,524	9,02E-07	=	MMP7
10	-3,522	0,002380527	=	ACTA2
19	-3,518	6,65E-05	=	LIG1
1	-3,511	0,000579731	=	LAMC1
1	-3,508	1,11E-06	j	SRSF11
8	-3,504	2,15E-06	j	KCNB2
21	-3,503	1,07E-06	x	LINC00113
1	-3,501	0,002056453	j	TEKT2
6	-3,498	0,005430124	n	CRISP2
11	-3,496	1,11E-06	=	TMEM123
19	-3,491	0,002214065	=	ILF3
3	-3,483	0,008085384	n	TTLL3
7	-3,481	1,60E-06	x	AC005024.1
22	-3,479	0,001324137	j	TXNRD2
7	-3,462	0,00173583	n	DPY19L2P1
2	-3,453	0,003596991	=	SCN1A
3	-3,449	3,35E-06	u	NA
5	-3,436	0,000237564	x	CTD-2201E18.4
19	-3,431	0,000261824	j	PLIN5
2	-3,423	0,002857113	j	IL1R1
21	-3,410	2,17E-06	n	AL035610.2
17	-3,409	2,62E-06	=	KCNJ16
11	-3,407	0,003097997	=	SOX6
22	-3,401	1,25E-05	j	DDX17
5	-3,399	2,55E-06	u	NA
2	-3,386	2,28E-05	n	IL1R1
1	-3,382	4,03E-05	=	LGALS8
11	-3,372	0,004344955	=	CREBZF
18	-3,371	0,001895388	=	AQP4-AS1
19	-3,371	0,000563306	=	KCNC3
15	-3,369	4,21E-06	=	TPM1
15	-3,368	0,000134098	=	TMC3
12	-3,354	0,000154823	j	R3HDM2
1	-3,347	1,04E-05	=	KIAA1324
14	-3,345	0,00010207	m	FRMD6-AS2
16	-3,339	9,79E-06	n	SMG1
10	-3,332	6,64E-05	j	ACSL5

16	-3,328	0,000335099	=	CDH16
16	-3,324	1,42E-05	j	METTL9
11	-3,318	0,00805006	j	SSSCA1
1	-3,307	5,44E-06	x	CDKN2C
11	-3,303	0,000128135	=	CD44
12	-3,299	0,003490081	j	HNF1A-AS1
2	-3,291	0,007090746	=	SCN1A
11	-3,273	6,30E-06	k	RP11-688I9.4
10	-3,269	0,00787701	j	ZFAND4
1	-3,269	6,27E-06	u	NA
8	-3,263	2,54E-05	j	TEX15
22	-3,257	4,70E-05	j	KIAA1671
12	-3,254	0,000106964	=	NCAPD2
12	-3,252	0,001083285	=	CLEC12A
2	-3,251	0,008745081	=	R3HDM1
2	-3,229	0,000378347	u	NA
11	-3,221	8,49E-06	=	RPS3
4	-3,217	0,009484555	=	FGG
21	-3,204	0,008482345	n	UMODL1
10	-3,192	4,61E-05	=	SORCS1
17	-3,183	0,000131924	j	MYO1C
9	-3,180	6,91E-05	k	CBWD5
22	-3,178	4,16E-05	=	PACSIN2
5	-3,177	1,58E-05	u	NA
9	-3,173	1,84E-05	=	GDA
8	-3,165	1,70E-05	k	AC083843.1
X	-3,165	5,49E-05	u	NA
16	-3,165	0,007485939	j	STX4
8	-3,160	1,77E-05	=	NDRG1
15	-3,159	0,009733359	=	CLPX
3	-3,156	1,86E-05	u	NA
3	-3,153	9,12E-05	=	CCNL1
6	-3,143	0,003275297	j	EZR
10	-3,141	0,006867731	j	SPAG6
11	-3,132	3,63E-05	j	NADSYN1
14	-3,125	0,001196755	j	PLEKHH1
7	-3,121	0,002012168	j	MAGI2-IT1
15	-3,104	3,86E-05	=	FAM174B
4	-3,104	0,000572021	j	ARAP2
X	-3,097	0,000145564	x	RP6-218J18.2
5	-3,094	0,000198588	m	COMMD10
19	-3,093	0,000644739	n	DMKN
1	-3,085	0,001817236	u	NA
9	-3,078	0,000763698	=	COL27A1
7	-3,076	0,000137995	n	RP5-1121A15.3
10	-3,070	0,000108782	=	SEMA4G
10	-3,066	5,12E-05	m	GSTO2

12	-3,063	0,000227714	x	LRCOL1
2	-3,055	2,83E-05	=	IGKC
6	-3,053	0,004337597	=	DDR1
16	-3,051	4,22E-05	=	C16orf89
19	-3,042	0,000105453	j	PTPRS
3	-3,042	4,03E-05	x	ACTR3P3
16	-3,039	6,45E-05	u	NA
3	-3,039	0,006448257	u	NA
14	-3,030	0,000939646	=	C14orf105
1	-3,029	0,000663109	n	TMEM234
19	-3,024	0,008526181	u	NA
3	-3,017	6,17E-05	u	NA
5	-3,015	7,64E-05	u	NA
14	-3,015	0,000118997	m	NOP9
2	-3,013	0,000355596	=	IHH
5	-3,013	0,006626795	=	FLT4
20	-3,007	0,00020816	n	RP5-1016A21.1
2	-3,006	4,01E-05	j	AC078941.1
2	-3,005	0,000171416	j	AC108938.5
16	-3,001	0,007025983	=	CES4A
1	-2,999	0,003454867	=	RP4-665N4.4
16	-2,998	9,04E-05	u	NA
10	-2,993	0,000748689	=	PFKFB3
12	-2,992	0,000542566	=	LDHB
13	-2,990	4,45E-05	k	RP11-124N19.3
5	-2,989	0,002824737	j	KDM3B
3	-2,988	0,002283552	=	CAMP
21	-2,982	0,000645481	m	CH507-42P11.8
1	-2,982	0,000146806	=	KAZN
15	-2,980	0,000414225	=	RP11-326L17.1
1	-2,971	6,04E-05	n	ERICH3
18	-2,971	0,006109993	=	FHOD3
1	-2,968	0,000593972	n	RP11-421L21.3
10	-2,967	0,001208093	j	ARMC3
4	-2,964	0,000342337	n	FBXL5
1	-2,963	0,003181097	=	ECHDC2
19	-2,960	8,01E-05	m	ITPKC
14	-2,953	0,000495463	k	MMP14
8	-2,948	0,000113404	k	ZNF252P
10	-2,944	0,001263014	j	MKX
11	-2,943	0,000305868	j	C11orf49
19	-2,938	0,000225836	=	ARHGEF1
14	-2,934	0,000311491	j	PPP4R4
2	-2,932	0,000373739	n	SPOPL
5	-2,931	0,000798941	u	NA
9	-2,926	0,001236937	n	KIF27
17	-2,923	0,000277215	k	BZRAP1

16	-2,921	0,006867731	=	TCF25
5	-2,921	0,000631681	=	TMEM167A
1	-2,918	7,93E-05	j	KIAA1324
1	-2,916	0,006614098	m	C1orf168
19	-2,916	0,002022063	u	NA
2	-2,913	0,00290713	j	HSPE1-MOB4
1	-2,910	8,89E-05	=	KIAA1324
14	-2,907	0,00012223	n	C14orf105
14	-2,907	0,000658623	j	NDRG2
18	-2,906	0,000101457	=	CHST9
12	-2,905	0,000151798	u	NA
3	-2,903	0,001284968	=	SHQ1
1	-2,900	0,001882438	j	RP1-35C21.1
10	-2,897	0,001217027	j	SORCS1
21	-2,896	0,00219649	j	TMEM50B
16	-2,886	0,000127566	j	RP11-58C22.1
11	-2,883	0,00012882	m	PHOX2A
5	-2,879	0,002840828	j	FAM153C
1	-2,877	0,006189716	j	HHAT
18	-2,877	0,001015156	=	SYT4
11	-2,876	0,00017101	=	CRACR2B
17	-2,873	0,000292075	x	TIAF1
15	-2,868	0,001404577	j	SCAND2P
1	-2,868	0,000129933	=	PKP1
15	-2,860	0,000108584	=	CRABP1
2	-2,844	0,000126688	u	NA
19	-2,843	0,000287791	j	ZNF544
15	-2,836	0,000692798	j	SIN3A
8	-2,833	0,000130723	n	NRG1
2	-2,819	0,000203909	k	AGFG1
20	-2,816	0,000987928	=	SRC
11	-2,816	0,000440612	j	UVRAG
19	-2,816	0,00427301	=	PRKD2
21	-2,812	0,000148304	n	AL035610.2
3	-2,812	0,000210499	u	NA
11	-2,806	0,003741583	=	PDZD3
1	-2,803	0,000213056	=	KIAA1324
3	-2,798	0,000200416	=	PLA1A
1	-2,798	0,002481582	n	PDE4DIP
11	-2,797	0,000245906	n	DOC2GP
10	-2,785	0,000506344	n	HECTD2
18	-2,784	0,002390572	j	DLGAP1
3	-2,781	0,000246466	n	RBM6
X	-2,778	0,000210499	=	BEX1
20	-2,767	0,000210499	x	RNU6-929P
1	-2,763	0,000433017	=	PLEKHG5
9	-2,753	0,003329736	=	CBWD3

3	-2,750	0,000471733	y	RP11-200A1.1
18	-2,748	0,001249546	=	ANKRD30B
7	-2,747	0,000557302	=	AC004540.4
19	-2,745	0,001607357	n	MBOAT7
1	-2,742	0,00560396	j	C1orf228
2	-2,736	0,000269984	=	AC079112.1
17	-2,735	0,000726606	k	RP11-219A15.1
8	-2,731	0,000591894	=	NCALD
17	-2,718	0,0027211	=	LRRC37A3
8	-2,711	0,000297956	x	RP11-21C17.1
2	-2,710	0,000670429	j	SCN1A
2	-2,708	0,000336507	u	NA
11	-2,702	0,000336564	=	MMP12
3	-2,701	0,001230094	=	DGKG
18	-2,701	0,003293799	=	SLC14A1
17	-2,697	0,001165139	m	MYBBP1A
2	-2,695	0,001181262	j	PKP4
1	-2,694	0,000460498	x	GPBP1L1
19	-2,689	0,001748849	=	ZNF497
2	-2,688	0,000366223	m	HOXD11
3	-2,685	0,000361389	=	SST
9	-2,683	0,003222421	x	RAPGEF1
11	-2,681	0,000407612	k	FZD4
14	-2,676	0,002793367	=	TTLL5
4	-2,672	0,000430268	u	NA
4	-2,671	0,005667991	j	ARHGEF38
1	-2,669	0,000642976	n	RAD54L
18	-2,667	0,000934938	j	CHST9
10	-2,667	0,000696622	=	EMX2OS
19	-2,665	0,000856516	k	CTD-2245F17.3
10	-2,661	0,00044513	x	GHITM
6	-2,659	0,000508566	m	MAP3K7
1	-2,658	0,000435426	=	KIAA1324
12	-2,654	0,004969331	j	SYT10
14	-2,654	0,000511102	x	SLC38A6
21	-2,654	0,001682432	n	C21orf58
X	-2,644	0,000546289	=	SYTL5
6	-2,641	0,003741583	=	CRISP3
11	-2,640	0,000765183	u	NA
12	-2,637	0,000663561	=	NUDT4
16	-2,636	0,001846386	m	CDH16
10	-2,636	0,000544901	u	NA
8	-2,636	0,000959409	x	RNU6-442P
1	-2,635	0,003689665	=	ESRRG
1	-2,633	0,003811725	m	TXLNA
11	-2,632	0,000624179	=	LDHA
11	-2,631	0,000541695	m	RNF169

5	-2,631	0,002402469	=	NIPBL
6	-2,630	0,00259764	=	STK19
19	-2,625	0,000616571	k	CTD-2017D11.1
12	-2,624	0,000658623	o	RP1-288H2.2
3	-2,623	0,001061491	=	FAM107A
3	-2,614	0,000594802	j	KBTBD12
15	-2,614	0,001131558	=	C2CD4B
14	-2,613	0,006626795	j	LRRC9
10	-2,611	0,000726606	=	MYO3A
7	-2,611	0,001067455	j	AC004540.4
2	-2,609	0,000899814	=	AGAP1
2	-2,607	0,000652336	u	NA
9	-2,600	0,005969245	j	TMC1
10	-2,599	0,000625958	u	NA
14	-2,595	0,002008316	n	HNRNPC
1	-2,590	0,009074138	j	RP11-413P11.1
8	-2,585	0,000711348	o	CASC8
2	-2,582	0,000742106	j	AC007389.3
3	-2,582	0,001141302	=	HMCES
21	-2,578	0,001200198	=	B3GALT5
9	-2,578	0,002432392	u	NA
22	-2,578	0,003313376	=	RBFOX2
11	-2,572	0,002113902	k	FNBP4
19	-2,569	0,000935	j	DNMT1
2	-2,566	0,001259259	n	RP11-310N16.1
2	-2,563	0,002323196	y	RP11-1223D19.1
15	-2,563	0,001141194	j	TICRR
10	-2,562	0,001121019	u	NA
4	-2,562	0,001014323	n	RBPJ
16	-2,560	0,002269817	j	ARHGAP17
2	-2,556	0,000858812	u	NA
8	-2,553	0,001093538	x	RP11-21C17.1
16	-2,549	0,000962053	=	PRSS21
10	-2,540	0,001692414	j	ADD3
4	-2,538	0,003489677	u	NA
16	-2,537	0,002088655	j	CPNE7
2	-2,535	0,001790952	j	RP11-310N16.1
11	-2,530	0,00259764	j	NUMA1
15	-2,524	0,00099861	j	GOLGA6L9
1	-2,522	0,005130348	=	PABPC4
7	-2,519	0,001151393	n	AC004540.4
1	-2,512	0,003518642	n	TMEM234
12	-2,510	0,001266722	m	IPO8
16	-2,507	0,00565365	=	USP31
17	-2,506	0,002224882	j	SPECC1
14	-2,504	0,003271054	=	MIS18BP1
X	-2,500	0,001131774	k	RP5-1158E12.3

2	-2,499	0,002362381	j	CFAP221
2	-2,498	0,00259709	j	RP11-287D1.3
10	-2,498	0,002438918	j	DDX50
1	-2,497	0,001919648	u	NA
11	-2,494	0,007705562	m	FNBP4
15	-2,490	0,001564639	n	UNC45A
11	-2,490	0,001492832	=	KLHL35
11	-2,489	0,001404577	k	RNF169
3	-2,485	0,001229936	=	IMPG2
1	-2,485	0,002168312	=	ANKRD35
12	-2,481	0,001250638	=	MYL6
12	-2,481	0,003600438	j	LINC00937
18	-2,480	0,005211929	j	OSBPL1A
11	-2,474	0,003866961	n	CRACR2B
19	-2,461	0,003907774	j	LDLR
6	-2,461	0,001895388	=	EPHA7
8	-2,460	0,002838313	k	UBXN8
11	-2,460	0,001533886	k	RNF169
15	-2,453	0,002581434	=	PRC1
1	-2,451	0,004634286	=	TRIM45
2	-2,448	0,00163042	n	AC108938.5
2	-2,443	0,003181097	n	CYP4F30P
9	-2,442	0,006279762	j	RP11-54D18.2
19	-2,441	0,003506232	=	SMIM24
15	-2,434	0,007557152	j	ADPGK
10	-2,431	0,001846747	u	NA
16	-2,428	0,006631043	j	HERPUD1
16	-2,425	0,004506049	=	PAM16
19	-2,425	0,002529508	j	CTXN1
1	-2,423	0,002096503	=	KIAA1324
17	-2,418	0,002126498	m	NOTUM
15	-2,415	0,007483262	n	PLCB2
6	-2,415	0,006818877	=	ENPP3
14	-2,414	0,007178081	n	HIF1A-AS2
17	-2,412	0,005032545	m	PTRH2
7	-2,408	0,008005937	j	DPY19L1P2
3	-2,405	0,002137754	j	C3orf58
12	-2,405	0,003179777	=	PPHLN1
X	-2,401	0,00259764	u	NA
18	-2,397	0,002134928	o	ZBTB7C
10	-2,393	0,002320126	u	NA
17	-2,391	0,004730883	=	SDK2
17	-2,389	0,005098378	=	CA4
11	-2,385	0,002688597	j	ASRGL1
21	-2,383	0,002381119	j	SIM2
1	-2,382	0,007876242	=	KIAA1324
2	-2,381	0,002705132	j	STAT1

11	-2,377	0,002383788	=	SYTL2
2	-2,373	0,002753343	u	NA
1	-2,372	0,004576323	=	KIAA1324
22	-2,368	0,002629825	=	IGLV3-1
1	-2,366	0,002719698	u	NA
11	-2,366	0,002720987	=	SYTL2
14	-2,364	0,003659688	n	PIGH
6	-2,354	0,009488381	j	CRISP2
18	-2,354	0,003006827	=	CHST9
12	-2,353	0,003282142	x	CTD-2021H9.2
3	-2,351	0,008085384	=	SPCS1
14	-2,350	0,002954186	=	IGHA2
2	-2,346	0,009006771	j	PCBP1-AS1
15	-2,343	0,003752121	k	LINS
20	-2,342	0,004001363	=	PLCB1
3	-2,341	0,004419192	k	RP11-297K7.1
8	-2,338	0,004337597	=	KB-1507C5.2
16	-2,335	0,008408248	=	DYNC1LI2
1	-2,334	0,005481908	j	GON4L
14	-2,333	0,004157568	u	NA
14	-2,330	0,003193873	=	ZFP36L1
11	-2,330	0,003174993	j	RNF169
2	-2,330	0,003202472	u	NA
2	-2,328	0,00356238	u	NA
9	-2,327	0,003233332	=	SET
2	-2,327	0,009873733	j	PSME4
21	-2,324	0,009967151	=	EVA1C
14	-2,321	0,003557617	j	CEP128
1	-2,317	0,004890047	=	SNAP47
6	-2,315	0,008053338	=	C6orf223
2	-2,315	0,00900831	m	FAM98A
12	-2,314	0,003893659	k	MPHOSPH9
11	-2,310	0,004587262	=	CTNND1
3	-2,307	0,003720342	u	NA
14	-2,305	0,005018107	j	SLC25A29
19	-2,303	0,004248398	=	RPSAP58
1	-2,303	0,008135517	x	ABCA4
15	-2,294	0,006885159	k	WDR73
3	-2,294	0,005832481	=	VEPH1
12	-2,288	0,005289906	j	MPHOSPH9
8	-2,286	0,009158691	j	NDUFAF6
1	-2,286	0,006284452	j	RP4-784A16.5
4	-2,282	0,004626352	k	CLDN22
22	-2,282	0,004248398	j	IGLL5
22	-2,281	0,005381618	j	PLA2G6
3	-2,279	0,005222156	j	KIF9-AS1
3	-2,278	0,004480034	=	VEPH1

17	-2,277	0,008078295	j	MAP3K3
18	-2,271	0,00839209	j	DLGAP1
3	-2,270	0,007123117	=	NDUFB5
2	-2,270	0,00437366	=	IGKV4-1
11	-2,270	0,004532548	u	NA
3	-2,269	0,005964815	=	NLGN1
12	-2,266	0,006090917	j	BCL2L14
17	-2,266	0,005285006	=	PIPOX
17	-2,265	0,007284389	j	BRIP1
3	-2,264	0,005540736	j	EGFEM1P
6	-2,263	0,004906301	=	HLA-A
20	-2,260	0,004898865	j	MIR646HG
12	-2,259	0,005318362	n	KNTC1
20	-2,258	0,005380857	=	FOXA2
15	-2,257	0,008653428	x	CTD-2240J17.2
14	-2,257	0,005983731	j	SLC25A29
5	-2,256	0,005627262	=	PPP2R2B
8	-2,253	0,005144023	j	COL14A1
12	-2,252	0,005756042	j	UBC
14	-2,252	0,007025983	o	RP11-638I2.6
3	-2,251	0,00490688	x	NLGN1-AS1
11	-2,248	0,005104998	x	RNU4ATAC5P
10	-2,247	0,009072344	j	CTNNA3
2	-2,247	0,00560396	=	M1AP
14	-2,247	0,006445841	u	NA
1	-2,247	0,005420769	k	FAM96AP2
3	-2,246	0,006774158	=	VEPH1
3	-2,243	0,009107896	=	KIAA1407
1	-2,242	0,006925131	=	CPSF3L
17	-2,240	0,006150932	=	CCL13
3	-2,240	0,0051666	=	CMTM6
1	-2,236	0,005567829	n	ERICH3
21	-2,232	0,005657829	=	PCP4
5	-2,232	0,008053338	x	FBXL7
3	-2,226	0,006454389	=	PARP14
8	-2,225	0,00907625	j	KB-1507C5.2
16	-2,225	0,007402665	=	BCO1
11	-2,224	0,006131164	=	KLHL35
3	-2,224	0,006300807	=	IL17RB
9	-2,224	0,006677393	k	CLCN3P1
2	-2,220	0,006129738	u	NA
1	-2,220	0,006614098	k	MORN1
12	-2,218	0,006222113	k	AC078864.1
11	-2,217	0,00597844	u	NA
1	-2,216	0,008768517	=	ERICH3
10	-2,215	0,007835019	=	RASGEF1A
10	-2,213	0,00906797	=	ZNF33A

15	-2,210	0,006129738	j	GOLGA6L9
5	-2,209	0,006129738	=	PIK3R1
10	-2,198	0,008152859	j	BCCIP
12	-2,197	0,008114608	=	SCNN1A
11	-2,197	0,007417777	=	NUMA1
14	-2,195	0,007772362	u	NA
12	-2,194	0,009334505	j	TPI1
4	-2,194	0,006624778	=	CXCL9
3	-2,193	0,007832036	=	SPTSSB
7	-2,192	0,007507256	u	NA
17	-2,192	0,00787701	m	SRSF1
17	-2,189	0,008353415	=	MPRIP
1	-2,186	0,007272427	k	OPN3
19	-2,185	0,008437065	=	SMIM7
3	-2,184	0,007531375	=	KBTBD12
2	-2,183	0,008289627	m	AC010883.5
6	-2,179	0,009059136	=	C6orf223
19	-2,178	0,009388325	x	GPR108
11	-2,177	0,00734157	u	NA
2	-2,176	0,008895158	j	AC007389.3
3	-2,167	0,008482345	=	GNB4
12	-2,164	0,007931932	=	PARPBP
20	-2,161	0,008838388	j	MIR646HG
11	-2,154	0,008792841	n	SNHG1
11	-2,153	0,009016427	=	TPCN2
1	-2,149	0,009382292	j	RPAP2
6	-2,149	0,00852642	=	RP11-632C17__A.1
14	-2,147	0,00955464	=	ACOT1
2	-2,142	0,009107896	j	RP11-418H16.1
11	-2,128	0,009382292	=	SCGB1D2
3	-2,127	0,009915783	=	GPR160
6	2,116	0,009978271	j	RUNX2
11	2,119	0,009915783	=	H19
5	2,122	0,009603345	=	VCAN
19	2,125	0,009637741	=	COL5A3
6	2,127	0,00942072	=	TPBG
15	2,128	0,009388325	j	LRRC28
2	2,130	0,009334505	o	AC012593.1
10	2,132	0,009439322	=	FGFR2
1	2,133	0,009603345	x	KIF26B
1	2,135	0,009074138	=	KIF26B
10	2,135	0,009766729	n	ADAMTS14
14	2,141	0,009260116	=	ZNF219
8	2,142	0,008804961	=	ZFHX4
10	2,144	0,009563551	j	CYP2E1
5	2,146	0,009490812	n	NKD2
2	2,147	0,008890567	u	NA

10	2,153	0,008244882	x	RNLS
12	2,153	0,008205491	=	MFAP5
8	2,154	0,008804961	=	RPL7
2	2,154	0,0083374	=	SCG2
22	2,156	0,008081401	k	LRP5L
22	2,157	0,009434768	n	ZDHHC8P1
2	2,158	0,008653428	u	NA
1	2,159	0,00924502	=	PTGER3
8	2,160	0,008016199	j	LINC00534
6	2,164	0,007892727	=	COL12A1
1	2,168	0,007835019	m	RP5-1065P14.2
1	2,171	0,007835019	j	MDM4
15	2,174	0,007537956	=	ALDH1A3
22	2,174	0,009157094	n	GUSBP11
9	2,174	0,008482345	j	COL27A1
5	2,177	0,007413863	n	RP11-348J24.2
10	2,177	0,009400258	=	CFAP70
2	2,179	0,007382158	u	NA
5	2,180	0,007133363	=	LIX1
21	2,185	0,008879785	j	NCAM2
2	2,185	0,007483262	x	RPL21P32
5	2,189	0,006770577	=	ADAMTS12
1	2,191	0,009302281	u	NA
2	2,192	0,009107896	u	NA
14	2,193	0,009325389	u	NA
11	2,195	0,006666946	=	MUC5B
11	2,196	0,007922045	j	ZBTB16
19	2,197	0,009617488	u	NA
7	2,201	0,006580809	j	LRRC4
1	2,204	0,007137502	=	CSF3R
1	2,206	0,008437065	u	NA
19	2,206	0,007090746	=	CEBPA-AS1
5	2,206	0,007056494	u	NA
10	2,207	0,006958677	n	ADAMTS14
13	2,207	0,006189716	x	RP11-54H7.4
1	2,207	0,009772272	u	NA
5	2,208	0,006614098	=	LOX
14	2,210	0,007957774	=	TMEM229B
17	2,213	0,007279347	n	ARHGAP27
2	2,215	0,006454389	k	AC108938.5
10	2,216	0,006809927	=	GRID1
17	2,217	0,005835706	k	TVP23C
2	2,219	0,008081401	x	PAX3
11	2,220	0,007828252	=	PHLDB1
7	2,223	0,007931932	j	IMMP2L
1	2,223	0,00573238	x	RP11-62I21.1
12	2,223	0,007663377	j	TCP11L2

21	2,224	0,006312485	=	RUNX1
1	2,226	0,008806668	=	PRG4
19	2,228	0,005530158	j	MUC16
2	2,230	0,005567829	j	AC007392.3
5	2,231	0,005982119	n	ZNF300
17	2,232	0,007507256	=	GAS7
4	2,234	0,007400748	k	RP11-362F19.1
13	2,235	0,005723866	=	POSTN
19	2,236	0,006432906	j	CD22
8	2,239	0,005368693	u	NA
2	2,242	0,006507179	u	NA
17	2,244	0,006614098	j	EFTUD2
16	2,245	0,006580809	j	VWA3A
15	2,246	0,005032545	=	IGF1R
4	2,246	0,005688839	j	LPHN3
2	2,248	0,006437518	j	SH3RF3
3	2,251	0,006597454	n	FAM86HP
13	2,252	0,005022313	o	MYO16-AS1
19	2,253	0,00516743	j	DPY19L3
3	2,253	0,005329352	=	ACKR4
15	2,254	0,005059663	j	GOLGA8B
5	2,254	0,005304436	n	SLC27A6
19	2,256	0,008081401	=	RHPN2
11	2,257	0,006783376	j	MRVI1
8	2,258	0,00629969	j	LPL
11	2,260	0,005059663	n	RIC3
4	2,260	0,007232986	=	COL25A1
12	2,261	0,009637741	j	PTPRQ
10	2,261	0,006269079	m	FRMPD2
19	2,262	0,006614098	=	LINC00662
10	2,268	0,00810396	=	COL17A1
12	2,269	0,005852781	n	ATN1
11	2,273	0,004546114	x	IGF2
1	2,273	0,006752011	k	MTX1
4	2,274	0,004999397	x	AP1AR
15	2,276	0,004961547	=	LRRC28
12	2,277	0,006495154	=	SCNN1A
2	2,278	0,006444273	=	DYSF
1	2,283	0,006421887	x	MYSM1
5	2,285	0,006464293	=	NKD2
1	2,286	0,005723866	=	AP006222.2
13	2,287	0,005775408	=	SUPT20H
8	2,288	0,00937852	n	AP3M2
19	2,289	0,005574982	n	NFIC
1	2,289	0,004825919	=	RP1-79C4.4
3	2,290	0,004969331	j	ROBO2
2	2,294	0,005318362	u	NA

16	2,296	0,005832481	=	CDH11
8	2,297	0,00437366	j	RP11-30J20.1
7	2,300	0,005285006	o	TWIST1
13	2,301	0,004344955	n	N4BP2L2
3	2,301	0,003811725	j	WNT7A
19	2,301	0,007541951	=	FOSB
1	2,302	0,003689665	n	CSMD2
10	2,305	0,005018107	=	SCART1
2	2,306	0,003623213	=	AC118345.1
1	2,308	0,003823643	k	RP11-195C7.3
4	2,308	0,009637741	n	ARHGAP10
1	2,309	0,005101297	u	NA
2	2,312	0,00492484	j	PCBP1-AS1
4	2,313	0,003873554	k	GUCY1A3
19	2,314	0,003413726	j	CTC-459F4.3
1	2,314	0,003666611	j	AMY1A
5	2,315	0,0083374	j	SYNPO
5	2,317	0,004159331	=	PCDHGA6
2	2,317	0,003356168	u	NA
12	2,318	0,006408055	=	MDM1
7	2,319	0,005654287	x	CRHR2
2	2,320	0,003662202	j	PXDN
12	2,321	0,005145781	n	TPCN1
12	2,326	0,009777021	j	GOLGA3
2	2,327	0,003262987	n	AC118345.1
3	2,328	0,003174993	k	CD47
7	2,331	0,003170749	j	COL1A2
4	2,331	0,003123776	=	SFRP2
1	2,332	0,004194356	k	RP11-34P13.13
15	2,332	0,003104853	=	TTC23
16	2,332	0,005654931	=	MRPL28
11	2,333	0,004933534	j	LTBP3
5	2,333	0,004520417	=	PAM
17	2,333	0,003527862	j	TANC2
1	2,333	0,005631822	n	RP11-504P24.2
3	2,334	0,003311926	n	FBXL2
20	2,335	0,003344275	m	GMEB2
2	2,337	0,003234251	x	NAB1
16	2,339	0,005075527	=	CDH11
12	2,340	0,005078023	j	RIMBP2
10	2,340	0,003330469	=	ADAMTS14
12	2,341	0,003961949	m	RP5-944M2.3
19	2,345	0,002902608	x	CTC-459F4.7
22	2,347	0,004401396	u	NA
10	2,348	0,003097997	=	ANXA8L1
20	2,350	0,003006827	j	FAM182B
7	2,353	0,004093388	j	AC004538.3

22	2,354	0,003174993	k	MIAT
7	2,355	0,002954186	m	COL1A2
1	2,356	0,005644114	u	NA
11	2,358	0,002931799	=	GUCY1A2
13	2,360	0,006476248	j	RP11-307N16.6
18	2,361	0,002686481	j	ALPK2
3	2,361	0,002931799	=	ALS2CL
3	2,364	0,004753247	m	PHLDB2
19	2,364	0,00437366	j	NFIX
4	2,365	0,002817055	u	NA
17	2,365	0,003170749	j	CCDC57
3	2,366	0,003498618	j	CTDSPL
12	2,366	0,007451421	=	TUBA1A
2	2,367	0,002539536	u	NA
4	2,372	0,002658666	j	RAPGEF2
3	2,372	0,002534	j	ZMAT3
11	2,372	0,002691842	u	NA
22	2,374	0,004403761	u	NA
11	2,375	0,005212155	n	TRIM22
11	2,375	0,005414075	=	TNNT3
10	2,376	0,004010135	n	FRMPD2
10	2,378	0,002954186	=	CH17-360D5.2
11	2,378	0,005923398	=	PHRF1
9	2,379	0,005983731	=	SVEP1
12	2,389	0,002857113	m	PPP1CC
17	2,391	0,00437366	n	C17orf70
1	2,392	0,002464912	=	COL16A1
19	2,393	0,002342309	k	ZNF587B
12	2,394	0,005098378	=	RP11-81H14.1
1	2,396	0,002283552	j	PRRX1
1	2,396	0,002283552	=	AMY1B
2	2,397	0,009016427	j	NDUFAF7
20	2,400	0,003222421	m	RALGAPA2
19	2,400	0,002305872	j	COL5A3
4	2,401	0,002367247	j	UBE2D3
2	2,402	0,002271712	u	NA
12	2,403	0,006269079	=	RAD52
13	2,404	0,002065041	=	POSTN
4	2,406	0,002285147	=	LPHN3
22	2,407	0,002113902	=	MMP11
8	2,411	0,002172487	m	ZFHX4
10	2,414	0,002103511	j	EMX2OS
16	2,415	0,002495306	=	RPS2
19	2,415	0,003466232	=	ILF3
6	2,417	0,003279078	j	SNX9
9	2,417	0,002182146	n	GLIS3-AS1
22	2,420	0,003345445	u	NA

19	2,421	0,003613532	n	CIRBP
10	2,421	0,003174993	u	NA
14	2,424	0,005618752	j	LRFN5
11	2,424	0,005381618	s	RP11-819M15.2
12	2,427	0,002323196	j	PLXNC1
20	2,427	0,003565424	j	CASC20
15	2,427	0,002269817	=	LRRC28
20	2,433	0,002305872	j	ISM1
13	2,435	0,001802322	o	MYO16-AS1
6	2,437	0,001689386	=	COL12A1
18	2,439	0,00185637	=	ZNF521
2	2,440	0,001639117	u	NA
8	2,442	0,001639117	=	CLU
20	2,443	0,002009682	n	ADAM33
1	2,443	0,00230502	=	SELP
2	2,444	0,001603922	n	AC118345.1
1	2,445	0,001776413	j	CSMD2
6	2,446	0,004848004	=	SYNE1
19	2,449	0,001540786	j	MUC16
13	2,449	0,003569232	=	FAM216B
1	2,449	0,002360543	j	LYPLAL1
20	2,450	0,001764473	=	LRRN4
7	2,451	0,002606098	=	IQCE
17	2,455	0,001836339	=	PLXDC1
11	2,456	0,008114608	m	CNGA4
9	2,458	0,001960245	j	GLIS3-AS1
3	2,458	0,008209403	=	CDHR4
1	2,459	0,005430124	n	SPAG17
4	2,460	0,001955406	n	EXOC1
20	2,460	0,009135502	n	DZANK1
2	2,462	0,003235751	n	ZEB2
8	2,463	0,003567995	k	ATP6V1B2
6	2,468	0,003811725	=	PTCHD4
4	2,468	0,003174993	=	ANTXR2
4	2,469	0,003276207	n	TAPT1-AS1
5	2,469	0,00136083	j	ADAMTS12
20	2,469	0,00139027	=	PMEPA1
1	2,470	0,001599668	=	AK5
11	2,471	0,001341185	=	GUCY1A2
20	2,475	0,003732399	x	FERMT1
2	2,477	0,001528405	n	DNMT3A
19	2,480	0,001670097	n	PRKCSH
6	2,481	0,006150932	=	RNF8
1	2,481	0,001882438	n	RP11-195C7.3
20	2,481	0,002287673	j	SYS1-DBNDD2
17	2,482	0,004093388	k	SLC16A3
1	2,482	0,001549354	j	TTC39A

12	2,487	0,001549354	=	NR2C1
6	2,487	0,001987272	u	NA
19	2,488	0,004999397	=	DOCK6
3	2,488	0,003803537	j	AGTR1
3	2,489	0,003344275	j	RP11-933H2.4
20	2,491	0,00137945	j	ISM1
1	2,494	0,002101173	n	CSF3R
21	2,495	0,002283265	m	AP000962.2
6	2,495	0,001180901	k	Y_RNA
14	2,495	0,00363362	j	NUMB
20	2,498	0,002134144	j	CASC20
19	2,498	0,00452199	=	PLEKHJ1
3	2,499	0,004301709	=	FAM198A
10	2,499	0,006986415	=	ANXA8L1
18	2,500	0,001612219	=	PIEZO2
1	2,502	0,002151543	j	MORN1
19	2,502	0,001315595	j	ANKRD27
12	2,503	0,001099868	m	LUM
3	2,504	0,001099868	o	LSAMP
5	2,507	0,003596991	=	FGF1
2	2,508	0,00210735	x	AC092162.1
2	2,509	0,003907774	j	PSD4
16	2,512	0,002242165	n	CLCN7
17	2,516	0,003823643	k	PLEKHM1P
1	2,517	0,001284968	j	CSMD2
6	2,518	0,006664912	=	CD24
6	2,520	0,005300355	=	MYLIP
17	2,523	0,001151393	n	COL1A1
11	2,524	0,002073524	j	TSKU
13	2,526	0,001099868	=	TNFRSF19
14	2,527	0,001074363	=	LTBP2
4	2,527	0,001557028	k	RP11-241F15.10
11	2,531	0,002390572	j	OSBPL5
17	2,535	0,001202778	=	DNAH9
15	2,536	0,000941732	j	TTC23
1	2,536	0,009603934	m	PTGS2
2	2,538	0,000971947	o	TWIST2
11	2,540	0,000989783	=	MRVI1
3	2,540	0,006809927	j	ROBO2
13	2,544	0,007932377	o	GPC5-AS1
6	2,544	0,008437065	j	AIM1
19	2,545	0,003275692	j	ZNF234
3	2,546	0,003659688	x	AGTR1
1	2,547	0,00087832	=	KIF26B
20	2,548	0,00245786	m	CSRP2BP
6	2,548	0,000912898	y	RP11-46B11.2
14	2,553	0,000830719	=	DUXAP10

4	2,553	0,0051666	j	LARP1B
11	2,556	0,000800168	=	IGF2
2	2,557	0,000930381	j	ANKRD36B
21	2,560	0,007474628	=	CH507-396I9.6
10	2,561	0,001781832	j	ANXA8L1
11	2,562	0,000830085	=	IGF2
1	2,566	0,000798941	k	PTGER3
12	2,566	0,000748333	n	RAP1B
4	2,566	0,00075205	u	NA
10	2,567	0,006653658	k	CCDC186
15	2,569	0,000824665	j	TTC23
15	2,570	0,000770302	=	HSP90B2P
5	2,572	0,004320615	=	NKD2
10	2,573	0,000939646	n	MMS19
16	2,573	0,005654931	j	GPR56
1	2,574	0,000781704	j	AMY1A
11	2,575	0,000868946	m	RP11-817J15.2
21	2,576	0,002737332	x	TEKT4P2
22	2,579	0,003626796	=	MICAL3
1	2,579	0,001438644	=	CSMD2
13	2,579	0,001338244	m	LCP1
20	2,579	0,000765183	j	MIR646HG
14	2,581	0,002044547	=	TTC6
12	2,588	0,001955406	=	C12orf75
17	2,589	0,000851548	=	FAM20A
1	2,591	0,000956787	n	KIAA1614
1	2,594	0,00805006	=	CFAP74
2	2,597	0,000692798	=	INPP5D
1	2,599	0,00516743	n	LRRC71
13	2,600	0,000748689	=	DNAJC15
11	2,600	0,000798941	m	SF1
3	2,600	0,001029491	=	RASA2
2	2,602	0,000648993	n	ANKRD36B
6	2,603	0,000660831	x	RP3-495K2.3
2	2,603	0,00224832	n	TRIP12
15	2,605	0,000673674	=	LRRC28
17	2,606	0,000993585	n	HOXB3
17	2,606	0,007993807	k	USP43
4	2,606	0,000724758	=	HHIP
19	2,607	0,000589914	n	ZNF507
5	2,607	0,000650001	j	ADAMTS12
15	2,612	0,000583094	i	LRRC28
2	2,612	0,006958677	n	KIAA2012
7	2,614	0,001121607	=	WDR86
2	2,617	0,005430124	=	NFE2L2
19	2,619	0,000678926	j	AC022153.1
6	2,619	0,000725627	m	HLA-DQB1

10	2,621	0,002695444	n	LIPN
6	2,622	0,000748333	=	CNR1
14	2,625	0,000792634	m	NOP9
6	2,627	0,000538497	j	PTCHD4
1	2,627	0,000572021	j	RP1-45C12.1
11	2,634	0,001167172	j	OSBPL5
12	2,638	0,000546977	=	CLIP1
4	2,642	0,00493429	x	ALG1L7P
19	2,643	0,000749155	j	POU2F2
4	2,646	0,000589914	j	TENM3
2	2,646	0,000726606	j	SLC30A6
17	2,648	0,000637092	=	COL1A1
16	2,648	0,00516413	j	GPR56
6	2,652	0,0083374	j	EYA4
15	2,657	0,000639337	=	FBN1
1	2,659	0,00060453	u	NA
2	2,662	0,002754334	n	CATIP
6	2,662	0,00436419	u	NA
17	2,672	0,003649876	=	WNK4
14	2,674	0,000377228	k	LINC01296
12	2,675	0,005429332	j	TXNRD1
19	2,676	0,000506855	j	NOTCH3
10	2,677	0,000395836	j	CH17-360D5.2
20	2,678	0,000377228	=	LRRN4
4	2,679	0,001682432	j	RP11-241F15.1
2	2,680	0,000537574	=	SCRN3
15	2,681	0,001130882	j	FAM189A1
7	2,681	0,000487625	j	AC093627.7
3	2,684	0,008109836	j	HHLA2
20	2,685	0,00042269	j	MIR646HG
12	2,687	0,003111085	j	NCAPD2
7	2,688	0,004982192	u	NA
2	2,689	0,000571296	n	ZEB2
11	2,689	0,00035988	=	IGF2
7	2,692	0,001933225	j	UBN2
3	2,692	0,000851548	=	MECOM
15	2,696	0,000694993	j	CCDC33
X	2,697	0,000337851	x	RP13-213K19.1
10	2,698	0,000366223	n	ADAMTS14
11	2,698	0,001391226	n	MYRF
15	2,699	0,000384202	=	TTC23
1	2,699	0,002283151	j	WDR63
1	2,705	0,000349552	n	PSMD4
3	2,707	0,000633911	j	GOLGA4
3	2,710	0,000832165	=	CLASP2
16	2,710	0,006548823	k	FBRS
11	2,712	0,000387288	=	RIC3

12	2,714	0,000541695	j	LIMA1
19	2,723	0,000279573	j	ZNF91
12	2,724	0,000763698	=	RP11-1143G9.5
11	2,725	0,000377228	j	RP11-627G23.1
3	2,728	0,000987325	j	FAM157A
16	2,730	0,000799816	=	MAF
3	2,733	0,000829934	n	CNOT10-AS1
2	2,736	0,000356409	=	AC007392.3
11	2,736	0,000253273	n	RIC3
2	2,737	0,000362352	n	AC104809.3
6	2,739	0,000267589	n	MLLT4
1	2,742	0,002269817	u	NA
7	2,744	0,003261542	j	POLR2J4
19	2,748	0,002691842	j	EMR3
2	2,748	0,004194491	j	KANSL1L
17	2,749	0,00189254	j	DNAH9
20	2,749	0,000339009	u	NA
12	2,751	0,000574629	=	GALNT9
5	2,751	0,000226973	m	RP11-138J23.1
20	2,756	0,00021987	x	MACROD2
18	2,757	0,000522298	n	CTD-2008L17.2
7	2,761	0,000487124	=	TMEM130
1	2,763	0,001522513	j	LINC01341
11	2,770	0,000374141	j	AHNAK
4	2,772	0,000217884	=	TENM3
19	2,775	0,005300355	j	LILRB2
4	2,777	0,00075205	j	CHRNA9
12	2,786	0,000312384	=	RP11-637A17.2
16	2,787	0,002438918	n	VWA3A
18	2,789	0,000540055	j	FHOD3
15	2,790	0,000529973	=	IL16
19	2,793	0,001725036	j	GRAMD1A
11	2,794	0,003423059	=	RIN1
12	2,796	0,000221663	=	RP11-284H19.1
19	2,798	0,000583015	j	LILRA1
2	2,804	0,001347318	n	KIAA2012
11	2,805	0,007772362	x	RP11-755E23.3
22	2,806	0,000227714	u	NA
2	2,807	0,000164763	u	NA
15	2,809	0,000332861	j	SMAD6
8	2,814	0,000182548	j	RP11-30J20.1
10	2,815	0,000148304	n	CH17-360D5.2
3	2,816	0,007402665	j	SETD5
5	2,816	0,000149591	=	CXCL14
22	2,817	0,001766925	m	AC006547.14
11	2,820	0,000639337	n	RP13-726E6.2
2	2,820	0,00014054	=	COL3A1

2	2,827	0,000221663	m	TRMT61B
19	2,829	0,001366873	k	ZNF836
1	2,838	0,000126688	=	FLG
1	2,838	0,00046933	k	WASH7P
15	2,846	0,000121996	x	LRRC28
15	2,846	0,000171416	m	PML
10	2,851	0,000418916	j	ANXA8
1	2,851	0,005983731	n	CFAP74
11	2,851	0,000624618	j	POU2AF1
15	2,857	0,00013748	=	AC022819.3
2	2,858	0,006454389	j	MAP3K19
11	2,862	0,001637381	n	ELMOD1
16	2,863	0,00012046	u	NA
21	2,875	0,000171416	j	ITSN1
6	2,876	0,000116929	x	AL590731.1
6	2,877	0,000897939	u	NA
15	2,878	0,000126688	=	RP11-35O15.2
16	2,882	0,000100179	j	LINC00922
20	2,885	0,001170724	=	SIRPB2
16	2,886	0,000253672	n	PKD1P1
21	2,887	0,000253672	u	NA
16	2,891	0,002707177	j	HAGHL
6	2,897	0,005300355	u	NA
22	2,897	0,000446894	j	FAM19A5
17	2,899	0,000692798	j	GOSR1
16	2,899	0,001888026	j	RP11-77K12.8
15	2,900	8,94E-05	=	IGF1R
1	2,902	0,00028482	j	ZBTB7B
X	2,908	9,30E-05	u	NA
17	2,909	0,004748866	j	USP43
5	2,914	7,91E-05	=	MEGF10
10	2,916	7,77E-05	=	GFRA1
6	2,919	0,000255	x	RREB1
8	2,924	0,000504485	=	FABP4
5	2,924	7,61E-05	k	MEGF10
2	2,924	8,93E-05	=	DYNC1I2
17	2,927	0,00053665	m	MINK1
4	2,930	0,005059663	j	HAND2-AS1
13	2,930	0,005735733	u	NA
5	2,930	7,22E-05	x	SLC12A7
1	2,935	0,000908004	j	HIVEP3
16	2,938	0,000748819	j	AC010547.9
7	2,940	8,10E-05	u	NA
4	2,942	9,81E-05	j	COL25A1
1	2,946	0,000527857	u	NA
19	2,947	0,002275765	u	NA
4	2,951	0,00011593	u	NA

X	2,952	5,80E-05	u	NA
6	2,957	7,51E-05	j	PHACTR1
13	2,958	0,007883165	u	NA
19	2,959	0,001963526	j	SLC25A23
1	2,963	6,06E-05	=	CSMD2
17	2,965	9,83E-05	=	AATK
11	2,970	0,000240688	j	CPT1A
1	2,971	9,04E-05	j	CFAP45
15	2,977	5,49E-05	o	RP11-654A16.1
10	2,979	0,000217722	=	VWA2
15	2,981	0,000406043	=	WDR61
6	2,983	0,001945584	j	PCMT1
1	2,983	6,62E-05	j	CADM3
6	2,986	0,000146635	x	RP11-254A17.1
19	2,986	0,009431948	=	DNAAF3
5	2,989	7,15E-05	=	VCAN
13	2,997	0,000219536	u	NA
21	2,999	6,56E-05	=	COL6A2
2	3,001	0,000258022	=	IL1RL1
19	3,001	8,90E-05	n	ZNF780A
10	3,002	5,75E-05	=	ADAMTS14
11	3,002	0,006688196	j	ELMOD1
11	3,005	8,67E-05	=	NELL1
11	3,005	0,000251989	=	TMEM138
11	3,021	6,43E-05	u	NA
2	3,023	0,000221556	=	ANKMY1
20	3,035	0,000118988	u	NA
20	3,042	0,000256286	u	NA
2	3,045	5,57E-05	=	MAP4K4
4	3,045	0,003262987	x	RNF150
8	3,048	2,99E-05	=	PKHD1L1
15	3,052	3,79E-05	k	CHD2
13	3,053	0,000471313	u	NA
8	3,053	0,000763698	=	BLK
12	3,055	0,000245906	x	KIAA1551
12	3,061	0,008151659	n	PAN2
7	3,062	2,75E-05	=	COL1A2
2	3,062	7,74E-05	u	NA
19	3,066	3,01E-05	j	NUDT19
22	3,067	0,000142266	k	MICALL1
1	3,068	6,25E-05	=	LRP8
3	3,068	0,000101215	m	LIPH
1	3,073	3,32E-05	j	PAPPA2
1	3,078	0,001781832	j	RP11-122M14.1
19	3,083	0,002937112	=	STXBP2
1	3,083	2,99E-05	=	CHIT1
16	3,086	4,22E-05	j	RP11-256I9.3

17	3,087	0,00787701	j	DNAI2
10	3,088	0,001049798	j	LIPN
12	3,089	0,000347502	m	GPR133
10	3,089	2,99E-05	=	ANXA8
17	3,099	3,01E-05	j	SPECC1
7	3,105	2,00E-05	j	COL1A2
5	3,106	2,28E-05	j	RP11-259O2.1
12	3,106	0,005983731	n	CCDC42B
6	3,113	0,000234975	u	NA
12	3,119	0,000142634	j	MPHOSPH9
20	3,130	2,63E-05	u	NA
10	3,133	3,76E-05	=	ANXA8L1
7	3,137	0,000148304	j	TNRC18
17	3,139	0,000861103	=	SEC14L1
2	3,141	4,24E-05	j	CFLAR
16	3,144	6,45E-05	x	GRIN2A
4	3,145	3,75E-05	j	CBR4
1	3,146	1,50E-05	u	NA
5	3,149	2,04E-05	x	SLC12A7
12	3,150	2,17E-05	=	KRT81
1	3,154	0,002390572	=	ILDR2
20	3,155	2,55E-05	j	CASC20
12	3,156	7,10E-05	j	KDM2B
5	3,156	3,18E-05	k	RP11-524L6.4
2	3,158	1,67E-05	n	LYPD6B
1	3,161	0,005532685	=	CCDC17
20	3,162	4,90E-05	x	FERMT1
X	3,162	0,000821422	j	NHSL2
1	3,167	0,006933565	=	CFAP74
11	3,170	0,000269984	j	TNNT3
15	3,174	1,29E-05	j	TTC23
12	3,187	1,10E-05	=	CCT2
10	3,192	0,000103969	=	ANXA8
19	3,194	5,31E-05	x	MUC16
6	3,195	1,09E-05	x	RP3-495K2.3
4	3,195	0,007402665	j	FAM47E
7	3,198	2,61E-05	u	NA
4	3,198	2,39E-05	n	RP11-696N14.1
12	3,198	0,000645481	n	DAZAP2
21	3,198	5,49E-05	j	AGPAT3
1	3,201	0,007553616	j	BAI2
21	3,203	0,00014098	j	COL6A2
20	3,204	0,007772362	u	NA
6	3,206	0,003596991	n	MLLT4
17	3,207	6,63E-05	=	SERPINF2
21	3,207	0,000440612	=	KCNE1
20	3,209	0,006917054	j	PLCB1

19	3,209	4,45E-05	=	TTYH1
22	3,213	1,35E-05	n	MMP11
4	3,220	0,001506648	n	TRIM2
1	3,222	8,50E-06	=	AMY1A
7	3,224	0,000322145	=	IQUB
4	3,230	4,18E-05	j	BMPR1B
2	3,231	1,14E-05	u	NA
19	3,234	0,000172451	=	HDGFRP2
19	3,236	0,001388527	=	CACNG6
4	3,236	0,007835019	=	SPATA18
20	3,236	0,004329704	m	DDRGK1
17	3,241	0,001338244	y	FLJ45079
15	3,242	1,85E-05	k	LRRK1
2	3,247	9,94E-06	x	ST6GAL2
4	3,250	1,77E-05	n	HHIP
2	3,252	1,94E-05	u	NA
2	3,257	1,35E-05	u	NA
10	3,262	8,67E-05	k	CH17-360D5.2
13	3,265	0,003525695	m	UGGT2
2	3,265	0,006783376	=	CYP4F62P
1	3,266	0,002754334	n	CSF3R
7	3,268	0,001321172	j	ZNF398
14	3,272	0,000485447	=	SLC24A4
3	3,272	5,06E-05	j	CADM2
8	3,275	0,00131082	=	CSMD3
4	3,277	0,000857587	=	DTHD1
4	3,282	2,28E-05	k	CFAP97
17	3,282	7,49E-06	k	HOXB3
1	3,284	6,94E-06	u	NA
13	3,294	5,99E-06	n	ARHGEF7
15	3,297	2,21E-05	=	ZNF106
2	3,298	0,004602809	j	LIMS1
14	3,300	0,000110823	=	G2E3
16	3,304	8,22E-06	n	ITGAX
12	3,310	4,51E-06	=	CPSF6
4	3,320	0,004806208	j	TET2
15	3,322	4,51E-06	=	TTC23
21	3,326	5,66E-05	j	SETD4
20	3,331	1,11E-05	u	NA
15	3,332	4,56E-06	o	RP11-35O15.1
10	3,336	5,38E-06	j	CH17-360D5.2
2	3,336	2,28E-05	u	NA
1	3,336	2,19E-05	n	CSF3R
5	3,345	5,71E-06	j	NKD2
11	3,349	0,001092756	j	ELMOD1
17	3,351	0,002767772	=	RP11-334C17.5
2	3,351	0,000726311	=	TUBA4B

12	3,352	4,41E-05	j	CCDC62
1	3,352	0,000523661	j	PTPRC
11	3,355	2,78E-05	=	C11orf88
22	3,359	7,81E-06	n	A4GALT
15	3,364	3,44E-06	x	RP11-294C11.4
14	3,367	0,003585423	u	NA
16	3,370	1,69E-05	j	GAS8
16	3,376	0,005688839	j	IL4R
14	3,387	2,55E-06	k	LINC01296
19	3,392	0,001011737	j	CRLF1
4	3,395	0,000446894	j	MAPK10
19	3,395	5,59E-06	n	PTPRS
4	3,400	4,51E-06	j	SH3D19
21	3,401	3,29E-06	j	BAGE2
3	3,402	2,49E-06	k	WNT7A
20	3,411	3,44E-06	u	NA
2	3,416	6,11E-06	j	SFXN5
2	3,417	3,13E-06	j	AC159540.1
3	3,421	2,24E-06	=	LSAMP
2	3,429	0,004228577	x	ST6GAL2
6	3,431	3,64E-06	u	NA
2	3,439	7,68E-06	=	DYSF
15	3,439	0,004248398	j	CSPG4
20	3,444	2,99E-06	u	NA
21	3,456	0,002371435	=	CH507-396I9.6
14	3,458	7,25E-05	n	SYNE3
3	3,458	5,22E-05	m	CDHR4
14	3,459	0,000387789	o	RP11-356O9.2
1	3,466	0,004969331	k	C1orf87
16	3,467	0,002712237	=	MEFV
12	3,468	0,007474628	j	KLRC1
3	3,471	0,00028482	=	PRKRIRP2
19	3,475	7,28E-05	j	CD22
16	3,477	0,002600373	o	RP11-459F6.3
4	3,478	2,49E-06	n	FRAS1
4	3,484	3,07E-06	=	HHIP
15	3,488	1,29E-06	o	RP11-20G13.1
10	3,488	2,49E-06	=	CUBN
19	3,489	0,002606098	n	ZNF414
10	3,492	0,00231028	j	KNDC1
10	3,497	2,27E-05	=	ANXA8
4	3,498	9,02E-05	=	EPGN
22	3,498	1,21E-06	=	EIF3D
17	3,498	0,000312761	x	AKAP10
1	3,499	1,02E-05	j	DAB1
14	3,502	0,000296877	x	MIPOL1
1	3,503	0,000110357	y	RP11-543E8.1

17	3,504	3,85E-06	k	BZRAP1
4	3,507	3,57E-05	u	NA
15	3,507	1,59E-05	j	LINC00923
3	3,512	1,87E-06	n	COL6A6
20	3,518	1,04E-05	u	NA
5	3,519	0,00414397	=	SNCAIP
12	3,520	9,23E-05	j	CEP83
4	3,528	9,55E-05	=	SORBS2
19	3,531	0,002683587	j	PCAT19
1	3,539	5,45E-06	o	C1orf132
13	3,545	6,98E-05	=	ANKRD10
2	3,546	0,002113163	j	UBXN4
2	3,551	0,000151096	j	FBXO11
15	3,558	1,21E-06	=	IGF1R
20	3,561	5,94E-06	u	NA
16	3,562	4,70E-05	j	AC010547.9
2	3,562	1,67E-06	j	ADD2
4	3,563	0,000420131	u	NA
3	3,566	5,94E-06	u	NA
5	3,567	7,08E-07	=	VCAN
2	3,572	2,24E-06	u	NA
14	3,572	0,001637381	j	TC2N
6	3,573	0,004180041	j	MYCT1
12	3,574	0,000935	=	PZP
22	3,577	0,000568929	j	XXbac-B444P24.8
3	3,580	6,60E-06	=	ZBTB20
1	3,581	6,73E-07	=	CLIC4
17	3,582	6,84E-06	j	ST6GALNAC2
20	3,583	0,001068699	m	RBCK1
1	3,584	8,42E-07	y	RP11-543E8.1
11	3,589	5,42E-07	u	NA
12	3,593	0,008353415	j	LMNTD1
16	3,598	0,000889801	n	FAM92B
15	3,600	0,001212074	m	CORO2B
1	3,601	0,000994495	=	CCDC17
21	3,602	0,001822675	n	RSPH1
2	3,602	0,001078967	n	SLC11A1
14	3,607	0,000146972	o	RP11-7F17.4
10	3,608	3,50E-06	j	FRMPD2
6	3,610	0,000153413	u	NA
15	3,612	6,83E-06	=	MAP1A
1	3,614	1,66E-05	u	NA
21	3,615	6,38E-07	n	AL078471.5
11	3,623	2,04E-05	j	NARS2
12	3,627	4,03E-07	=	MDM2
15	3,627	5,37E-07	=	TTC23
15	3,629	0,002212564	=	ABHD17C

2	3,634	8,93E-06	=	GLS
11	3,636	0,000234975	j	SBF2
6	3,638	0,00354439	u	NA
14	3,639	0,001315595	j	LRFN5
6	3,651	0,000798941	x	RP1-6P5.2
2	3,651	3,70E-06	m	NR4A2
7	3,659	0,002080827	=	RSPH10B2
19	3,661	3,13E-05	=	CLEC17A
15	3,662	1,62E-06	m	NOX5
3	3,675	2,74E-05	=	GRAMD1C
15	3,684	2,85E-07	j	TTC23
3	3,692	0,001088463	j	MYH15
2	3,713	1,81E-06	j	CYP26B1
1	3,715	1,22E-06	n	TMEM234
5	3,721	0,001032819	j	TTC23L
20	3,734	5,36E-07	j	LRRN4
19	3,738	1,42E-06	=	TTYH1
7	3,749	0,00382441	=	PNPLA8
22	3,750	9,21E-07	u	NA
17	3,755	0,001067455	u	NA
16	3,755	2,21E-07	m	VWA3A
1	3,759	2,30E-06	=	SORT1
15	3,760	7,47E-07	k	CTD-2311M21.2
17	3,766	6,22E-06	j	GLOD4
14	3,768	0,000456952	=	STON2
3	3,769	2,52E-07	j	COL6A6
4	3,773	3,45E-06	u	NA
3	3,773	9,92E-07	=	ST6GAL1
15	3,776	1,23E-07	=	IGF1R
9	3,777	1,23E-07	u	NA
19	3,778	0,00060977	n	GRIN3B
13	3,795	0,000872422	k	ATP11A
22	3,796	2,15E-05	j	KIAA1671
8	3,798	4,27E-07	=	FABP4
10	3,805	4,17E-07	j	KCNMA1
11	3,811	0,000104782	j	TMEM9B-AS1
19	3,817	0,001196755	n	CAPN12
1	3,823	3,19E-05	j	TAL1
1	3,823	2,50E-05	x	C1orf140
18	3,835	2,49E-06	=	B4GALT6
6	3,846	8,07E-08	n	PTCHD4
11	3,846	8,02E-06	n	BEST1
2	3,847	1,36E-07	n	PKP4
4	3,852	0,00937852	=	C4orf22
12	3,859	9,12E-08	n	RP11-71J4.2
14	3,866	9,11E-05	j	AKAP6
3	3,871	1,10E-06	=	CADM2

1	3,880	0,000502005	j	CFAP57
21	3,884	0,000282078	n	C21orf58
15	3,890	5,05E-08	j	IGF1R
3	3,890	0,009594259	=	IQSEC1
20	3,894	1,15E-07	x	FERMT1
19	3,896	2,98E-06	=	FFAR1
14	3,899	4,52E-05	=	MAP3K9
15	3,907	6,67E-08	k	RP11-6O2.2
16	3,909	1,15E-07	j	SMG1P1
2	3,913	0,001391846	=	BIRC6
11	3,913	3,72E-06	n	FLI1
20	3,927	5,52E-06	j	PTPRT
20	3,928	1,66E-07	u	NA
18	3,936	0,000309541	j	ATP9B
10	3,937	4,52E-08	n	RP11-144G6.12
2	3,939	1,01E-07	k	COMMD1
11	3,945	0,000378347	=	RELA
8	3,951	0,000832165	=	UNC5D
19	3,955	1,19E-07	j	RGL3
16	3,960	5,91E-07	u	NA
17	3,963	0,000250098	=	LINC00854
7	3,974	0,001637381	=	RSPH10B
1	3,985	0,000179354	=	NBPF15
3	3,985	0,00043408	j	GOLIM4
1	3,998	0,000171416	j	CFAP57
2	4,000	1,19E-06	j	CATIP
1	4,006	3,04E-08	x	RP11-63B19.1
1	4,007	9,58E-05	m	SYTL1
6	4,010	0,001417651	=	ZNF451
7	4,013	3,44E-06	u	NA
2	4,014	4,69E-05	j	AC104809.3
5	4,014	0,0003661	j	RP11-1415C14.4
19	4,015	0,000349552	=	PPP1R13L
17	4,047	0,000176146	n	CRLF3
13	4,065	0,000129933	k	MRPS31
6	4,067	5,68E-08	u	NA
14	4,084	0,00011592	=	NPAS3
3	4,089	1,28E-05	u	NA
19	4,091	7,82E-07	n	TLE2
2	4,098	0,000366223	=	IL1R2
1	4,098	0,00014098	j	AMY1C
2	4,102	1,02E-08	j	LINC01473
12	4,139	6,66E-09	j	CPSF6
3	4,141	8,75E-09	m	COL6A6
7	4,143	7,07E-06	n	SMURF1
19	4,148	0,000448552	j	RAVER1
2	4,151	0,000150952	=	SLC4A10

1	4,164	7,51E-05	=	DENND4B
2	4,167	5,36E-05	=	CCDC108
17	4,169	0,000105453	n	DBF4B
7	4,179	3,72E-08	j	NAMPT
2	4,183	3,31E-05	=	KIAA2012
9	4,197	0,002580172	j	IFT74
2	4,203	4,15E-05	=	TNS1
20	4,211	1,66E-08	u	NA
12	4,220	1,34E-08	=	TSPAN9
3	4,223	7,89E-09	j	COL6A6
2	4,226	2,46E-06	j	KIAA2012
4	4,226	6,16E-09	=	RP11-362F19.1
6	4,227	0,000126688	=	ECT2L
4	4,246	3,24E-05	=	MAN2B2
19	4,250	5,71E-06	=	ZFP36
19	4,252	6,32E-05	=	ICAM4
15	4,252	2,86E-09	j	RP11-20G13.1
7	4,254	0,00054332	x	RP11-397J20.1
19	4,262	4,75E-05	j	CTD-2207O23.3
13	4,271	5,79E-09	k	GAS6-AS1
11	4,280	4,24E-05	j	LMO2
14	4,292	5,07E-05	j	ACIN1
18	4,304	3,71E-05	n	GATA6
20	4,305	8,45E-05	=	NNAT
20	4,307	2,14E-06	=	BTBD3
17	4,324	2,94E-05	=	MYL4
19	4,327	2,71E-05	j	GIPR
1	4,330	1,76E-06	=	ANGPTL7
1	4,340	2,71E-05	=	FHAD1
12	4,343	1,30E-09	n	RAP1B
7	4,352	0,005342271	k	STAG3L5P-PVRIG2P-PILRB
11	4,377	3,36E-05	=	MS4A1
17	4,379	2,70E-05	k	ZNF652
1	4,384	1,20E-08	j	TRIM11
5	4,389	2,62E-05	=	ICE1
3	4,398	4,43E-09	m	NISCH
1	4,403	2,85E-09	j	PAPPA2
4	4,421	1,53E-09	=	ADH1B
1	4,427	9,00E-10	u	NA
16	4,435	3,06E-05	m	ZFPM1
5	4,449	4,45E-05	n	CTD-2544H17.1
2	4,454	1,93E-05	j	MGAT5
3	4,483	1,40E-05	n	RBM6
11	4,489	1,16E-05	=	CFL1
16	4,490	0,008085384	=	ZNF598
12	4,495	8,66E-05	j	GALNT9
2	4,499	1,51E-05	x	PLCL1

2	4,512	3,51E-05	=	ATG4B
16	4,521	1,08E-07	x	ZNF263
4	4,522	6,67E-10	=	RP11-362F19.1
2	4,524	1,15E-05	=	WIPF1
12	4,527	1,11E-05	u	NA
6	4,528	3,43E-05	n	FAM46A
6	4,548	2,33E-10	o	PTCHD4
11	4,549	8,45E-06	m	BCL9L
3	4,563	1,24E-05	=	LINC00969
17	4,571	5,34E-06	=	ALOX15
12	4,574	8,77E-10	j	WNK1
9	4,578	5,66E-08	u	NA
3	4,579	1,37E-09	n	COL6A6
17	4,583	0,00573238	=	MYL4
2	4,589	2,36E-07	=	UBE2F
3	4,594	5,87E-06	n	LINC01267
12	4,598	8,23E-06	j	LTA4H
1	4,600	6,23E-06	j	SSBP3
16	4,621	5,48E-06	k	EIF3C
16	4,632	7,80E-06	m	STUB1
1	4,656	1,91E-09	=	ESRRG
10	4,686	2,67E-07	=	FRMPD2
1	4,691	4,34E-06	j	SELP
5	4,699	4,43E-08	=	TCOF1
3	4,704	2,15E-06	m	AC024560.3
15	4,706	8,99E-11	k	OR4N3P
20	4,717	9,62E-09	u	NA
11	4,745	2,07E-05	o	RP11-755E23.3
1	4,769	2,07E-06	j	ZMYM4
3	4,773	3,72E-06	=	ANKUB1
7	4,784	0,001217027	n	STAG3L5P-PVRIG2P-PILRB
16	4,785	1,87E-06	j	QPRT
10	4,791	5,30E-10	=	SCART1
18	4,793	4,25E-06	k	MBD1
19	4,795	2,40E-06	j	NFIC
16	4,801	0,001168283	=	ZNF23
19	4,802	1,18E-05	n	CAPN12
11	4,813	1,37E-09	=	SCGB1A1
3	4,828	2,98E-11	u	NA
12	4,830	1,83E-06	=	AACS
4	4,891	2,78E-11	u	NA
19	4,919	2,79E-05	j	PPP6R1
21	4,922	5,45E-07	n	C21orf58
6	4,927	4,17E-06	n	FAM120B
8	4,931	1,13E-05	n	LONRF1
3	4,931	1,13E-11	=	COL6A6
3	4,948	2,76E-11	j	COL6A6

9	4,948	0,005574982	j	GLIS3
12	4,951	1,53E-05	j	GPR133
16	4,965	1,09E-06	m	PLK1
5	4,995	1,35E-05	k	FAM153C
19	4,998	4,20E-07	j	CSNK1G2
2	5,022	2,45E-08	n	TUBA4B
14	5,033	3,53E-07	j	WDR20
1	5,068	1,70E-05	u	NA
19	5,072	1,17E-07	n	CAPN12
16	5,089	3,01E-12	u	NA
16	5,109	1,41E-07	n	FAM92B
15	5,111	4,12E-12	=	SPINT1
4	5,128	3,26E-06	j	NUDT6
7	5,139	7,28E-05	=	RSPH10B2
6	5,145	1,88E-12	m	PTCHD4
9	5,169	0,007876242	=	GLIS3
11	5,172	2,32E-07	u	NA
7	5,180	4,78E-06	j	CREB5
6	5,181	1,86E-12	u	NA
3	5,181	2,93E-12	j	COL6A6
19	5,192	1,52E-07	m	HMG20B
6	5,211	8,56E-13	u	NA
2	5,231	4,00E-07	=	NEU4
19	5,234	1,08E-07	j	GLTSCR2
22	5,239	7,43E-08	=	PITPNB
3	5,249	7,82E-08	j	ABHD6
20	5,265	8,83E-10	=	SOGA1
12	5,265	5,03E-08	j	ENO2
16	5,297	5,32E-08	j	NFATC3
3	5,299	1,44E-06	=	MUC4
2	5,300	7,21E-08	o	AC074011.2
5	5,301	6,17E-12	n	MTRR
1	5,304	1,11E-09	j	LEPR
14	5,304	1,40E-07	k	DDHD1
22	5,373	0,00805006	=	P2RX6
21	5,387	4,72E-08	j	KCNJ15
6	5,403	1,55E-13	x	RP11-254A17.1
15	5,420	0,000332162	=	NOX5
1	5,425	8,19E-11	=	ZBTB8B
1	5,426	1,94E-08	j	CDK11B
22	5,438	2,40E-06	n	XXbac-B444P24.8
17	5,442	1,04E-08	n	DNAH2
3	5,443	3,85E-08	n	GRAMD1C
7	5,445	1,88E-06	n	SUN1
2	5,446	7,61E-09	n	AC093838.4
3	5,455	2,20E-08	y	RP11-451B8.1
19	5,467	1,33E-08	=	HNRNPM

1	5,487	1,96E-10	j	HSPG2
13	5,489	2,87E-08	j	UPF3A
4	5,536	2,03E-08	=	NDNF
2	5,543	2,78E-12	x	ST6GAL2
12	5,547	2,15E-12	u	NA
6	5,575	4,20E-07	m	KLHL31
19	5,587	7,43E-09	j	WTIP
11	5,589	4,82E-09	j	USP2
7	5,592	4,23E-07	=	CAV2
7	5,593	1,40E-07	k	STAG3L4
2	5,609	5,42E-09	j	CFLAR
X	5,632	0,000242605	j	USP9X
3	5,642	3,93E-09	n	RP5-966M1.6
12	5,647	4,18E-08	=	CLEC1B
4	5,657	4,29E-09	=	FGF5
14	5,658	5,29E-09	=	KCNH5
6	5,681	1,90E-08	n	NFYA
10	5,701	1,60E-08	=	ANXA8
15	5,712	9,08E-09	=	PAK6
2	5,713	1,97E-09	=	RABL2A
5	5,726	2,55E-09	n	SPEF2
7	5,739	6,88E-15	m	RP11-328J2.1
13	5,769	4,52E-09	u	NA
17	5,774	1,06E-12	j	HOXB3
16	5,775	1,35E-09	k	FUS
10	5,787	2,85E-09	j	CPEB3
6	5,789	5,66E-15	x	RP11-254A17.1
3	5,802	5,54E-08	u	NA
7	5,831	3,43E-05	o	RP11-449P15.2
3	5,841	6,01E-10	=	MKRN2
2	5,842	9,00E-09	x	ALPPL2
16	5,865	9,21E-12	j	TERF2
12	5,909	2,96E-14	j	KCNA1
2	5,926	3,53E-10	j	GREB1
1	5,961	3,99E-09	=	GPBP1L1
8	5,978	6,61E-07	j	PRKDC
6	5,982	2,28E-09	=	KIF13A
3	5,983	5,09E-13	j	BCL6
14	5,988	1,63E-15	=	CHGA
11	5,992	6,16E-10	=	MAML2
7	5,996	6,14E-05	=	RBM33
2	6,001	1,02E-12	m	NR4A2
15	6,004	5,27E-10	=	SYNM
1	6,008	8,02E-16	=	AMY1B
12	6,025	1,06E-14	=	KCNA1
2	6,028	1,85E-10	=	ARMC9
19	6,031	2,29E-10	j	ZNF701

22	6,125	9,74E-14	=	DEPDC5
14	6,133	6,61E-11	=	ZFP36L1
16	6,139	1,20E-11	=	BAIAP3
20	6,144	2,52E-10	j	ITCH
21	6,149	1,32E-10	j	C2CD2
12	6,224	1,30E-12	u	NA
19	6,230	2,09E-11	=	R3HDM4
15	6,231	8,41E-11	j	MGA
3	6,293	7,23E-11	n	TNIK
14	6,312	3,97E-11	j	KTN1-AS1
22	6,384	1,92E-11	u	NA
2	6,481	1,03E-17	n	TEX41
3	6,487	1,36E-17	j	RP11-615J4.3
4	6,504	4,17E-12	n	FRYL
6	6,505	2,46E-10	j	MDN1
10	6,540	4,83E-12	j	MXI1
2	6,570	2,19E-12	n	IFT172
10	6,652	9,06E-13	k	SFXN2
8	6,670	2,46E-09	x	RPS3AP34
3	6,684	2,34E-12	k	ZBTB20
16	6,784	3,90E-13	n	DDX19A
22	6,784	3,17E-13	j	PPARA
8	6,789	6,63E-19	n	RALYL
2	6,791	5,09E-13	j	EPAS1
21	6,798	2,12E-13	j	PCNT
20	6,808	1,66E-13	=	PSMF1
17	6,832	9,83E-15	=	BCAS3
17	6,998	3,46E-14	=	GAS7
20	7,057	2,69E-14	j	RALY
12	7,171	1,18E-14	=	ANKRD13A
10	7,182	1,99E-14	j	GFRA1
7	7,272	1,62E-10	j	CTTNBP2
6	7,317	2,49E-14	m	CD83
15	7,424	3,19E-21	x	LRRC28
14	7,462	2,61E-15	j	CDKL1
20	7,519	8,91E-16	=	GZF1
11	7,524	7,79E-16	m	NXF1
14	7,634	2,61E-15	k	SOS2
12	7,698	1,56E-16	j	ERC1
2	7,759	1,63E-15	m	NR4A2
1	7,987	1,18E-17	=	PLEKHO1
16	8,183	1,23E-18	k	FUS
3	8,449	3,04E-19	n	RPUSD3
4	8,568	6,96E-19	j	CBR4
9	8,570	3,41E-12	j	XPA
3	8,691	1,07E-25	j	RP11-615J4.3
8	9,611	5,97E-18	=	ZBTB10

20	9,697	2,56E-23	u	NA
10	10,265	2,99E-25	n	RP11-144G6.12

7. References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. CA Cancer J Clin. 2019;69(1):7–34.

2. Cancer Facts & Figures 2019 | American Cancer Society [Internet]. [cited 2020 Feb 14]. Available from: https://www.cancer.org/research/cancer-facts-statistics/all-cancer-facts-figures/cancer-facts-figures-2019.html

3. Momenimovahed Z, Tiznobaik A, Taheri S, Salehiniya H. Ovarian cancer in the world: epidemiology and risk factors. Int J Womens Health. 2019 Apr 30;11:287–99.

4. Cancer Research UK. Ovarian cancer statistics. https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/ovarian-cancer/survival. 2020.

5. Reid BM, Permuth JB, Sellers TA. Epidemiology of ovarian cancer: a review. Cancer Biol Med. 2017 Feb;14(1):9–32.

6. Jacobs IJ, Menon U, Ryan A, Gentry-Maharaj A, Burnell M, Kalsi JK, et al. Ovarian cancer screening and mortality in the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS): a randomised controlled trial. The Lancet. 2016 Mar;387(10022):945–56.

7. van Nagell JR, Hoff JT. Transvaginal ultrasonography in ovarian cancer screening: current perspectives. Int J Womens Health. 2013 Dec 20;6:25–33.

8. Sölétormos G, Duffy MJ, Othman Abu Hassan S, Verheijen RHM, Tholander B, Bast RC, et al. Clinical Use of Cancer Biomarkers in Epithelial Ovarian Cancer. Int J Gynecol Cancer. 2016 Jan;26(1):43–51.

9. Lheureux S, Gourley C, Vergote I, Oza AM. Epithelial ovarian cancer. The Lancet. 2019 Mar 23;393(10177):1240–53.

10. Kuchenbaecker KB, Hopper JL, Barnes DR, Phillips K-A, Mooij TM, Roos-Blom M-J, et al. Risks of Breast, Ovarian, and Contralateral Breast Cancer for BRCA1 and BRCA2 Mutation Carriers. JAMA. 2017 Jun 20;317(23):2402–16.

11. Menon U, Karpinskyj C, Gentry-Maharaj A. Ovarian Cancer Prevention and Screening. Obstet Gynecol. 2018;131(5):909–27.

12. Penninkilampi R, Eslick GD. Perineal Talc Use and Ovarian Cancer: A Systematic Review and Meta-Analysis. Epidemiol Camb Mass. 2018;29(1):41–9.

13. Vaughan S, Coward JI, Bast Jr. RC, Berchuck A, Berek JS, Brenton JD, et al. Rethinking Ovarian Cancer: Recommendations for Improving Outcomes. Nat Rev Cancer. 2011 Sep 23;11(10):719–25.

14. Prat J, D'Angelo E, Espinosa I. Ovarian carcinomas: at least five different diseases with distinct histological features and molecular genetics. Hum Pathol. 2018 Oct;80:11–27.

15. Prat J. FIGO's staging classification for cancer of the ovary, fallopian tube, and peritoneum: abridged republication. J Gynecol Oncol. 2015 Apr;26(2):87–9.

16. Vang R, Shih I-M, Kurman RJ. OVARIAN LOW-GRADE AND HIGH-GRADE SEROUS CARCINOMA: Pathogenesis, Clinicopathologic and Molecular Biologic Features, and Diagnostic Problems. Adv Anat Pathol. 2009 Sep;16(5):267–82.

17. Kurman RJ, Shih I-M. The Dualistic Model of Ovarian Carcinogenesis. Am J Pathol. 2016 Apr;186(4):733–47.

18. Yemelyanova A, Vang R, Kshirsagar M, Lu D, Marks MA, Shih IM, et al. Immunohistochemical staining patterns of p53 can serve as a surrogate marker for TP53 mutations in ovarian carcinoma: an immunohistochemical and nucleotide sequencing analysis. Mod Pathol. 2011 Sep;24(9):1248–53.

19. The Cancer Genome Atlas Research Network. Integrated genomic analyses of ovarian carcinoma. Nature. 2011 Jun;474(7353):609–15.

20. Ahmed AA, Etemadmoghadam D, Temple J, Lynch AG, Riad M, Sharma R, et al. Driver mutations in TP53 are ubiquitous in high grade serous carcinoma of the ovary. J Pathol. 2010 May;221(1):49–56.

21. Konstantinopoulos PA, Ceccaldi R, Shapiro GI, D'Andrea AD. Homologous Recombination Deficiency: Exploiting the Fundamental Vulnerability of Ovarian Cancer. Cancer Discov. 2015 Nov;5(11):1137–54.

22. Ciriello G, Miller ML, Aksoy BA, Senbabaoglu Y, Schultz N, Sander C. Emerging landscape of oncogenic signatures across human cancers. Nat Genet. 2013 Oct;45(10):1127–33.

23. Ballabio S, Craparotta I, Paracchini L, Mannarino L, Corso S, Pezzotta MG, et al. Multisite analysis of high-grade serous epithelial ovarian cancers identifies genomic regions of focal and recurrent copy number alteration in 3q26.2 and 8q24.3. Int J Cancer. 2019 Nov 15;145(10):2670–81.

24. Tothill RW, Tinker AV, George J, Brown R, Fox SB, Lade S, et al. Novel Molecular Subtypes of Serous and Endometrioid Ovarian Cancer Linked to Clinical Outcome. Clin Cancer Res. 2008 Aug 15;14(16):5198–208.

25. Australian Pancreatic Cancer Genome Initiative, ICGC Breast Cancer Consortium, ICGC MMML-Seq Consortium, ICGC PedBrain, Alexandrov LB, Nik-Zainal S, et al. Signatures of mutational processes in human cancer. Nature. 2013 Aug;500(7463):415–21.

26. Macintyre G, Goranova TE, De Silva D, Ennis D, Piskorz AM, Eldridge M, et al. Copy number signatures and mutational processes in ovarian carcinoma. Nat Genet. 2018 Sep;50(9):1262–70.

27. Paracchini L, Mannarino L, Craparotta I, Romualdi C, Fruscio R, Grassi T, et al. Regional and temporal heterogeneity of epithelial ovarian cancer tumor biopsies: implications for therapeutic strategies. Oncotarget [Internet]. 2016 Jul 9 [cited 2020 Mar 18];5(0). Available from: http://www.oncotarget.com/index.php?journal=oncotarget&page=article&op=view&path%5B%5D=10505

28. Beltrame L, Di Marino M, Fruscio R, Calura E, Chapman B, Clivio L, et al. Profiling cancer gene mutations in longitudinal epithelial ovarian cancer biopsies by targeted next-generation sequencing: a retrospective study. Ann Oncol Off J Eur Soc Med Oncol. 2015 Jul;26(7):1363–71.

29. Lheureux S, Gourley C, Vergote I, Oza AM. Epithelial ovarian cancer. The Lancet. 2019 Mar 23;393(10177):1240–53.

30. Martín-Cameán M, Delgado-Sánchez E, Piñera A, Diestro MD, De Santiago J, Zapardiel I. The role of surgery in advanced epithelial ovarian cancer. ecancermedicalscience [Internet]. 2016 Aug 17 [cited 2020 Feb 24];10. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4990056/

31. du Bois A, Reuss A, Pujade-Lauraine E, Harter P, Ray-Coquard I, Pfisterer J. Role of surgical outcome as prognostic factor in advanced epithelial ovarian cancer: a combined exploratory analysis of 3 prospectively randomized phase 3 multicenter trials: by the Arbeitsgemeinschaft Gynaekologische Onkologie Studiengruppe Ovarialkarzinom (AGO-OVAR) and the Groupe d'Investigateurs Nationaux Pour les Etudes des Cancers de l'Ovaire (GINECO). Cancer. 2009 Mar 15;115(6):1234–44.

32. Lisio M-A, Fu L, Goyeneche A, Gao Z-H, Telleria C. High-Grade Serous Ovarian Cancer: Basic Sciences, Clinical and Therapeutic Standpoints. Int J Mol Sci. 2019 Feb 22;20(4).

33. Piccart MJ, Bertelsen K, Stuart G, Cassidy J, Mangioni C, Simonsen E, et al. Long-term follow-up confirms a survival advantage of the paclitaxel-cisplatin regimen over the cyclophosphamide-cisplatin combination in advanced ovarian cancer. Int J Gynecol Cancer Off J Int Gynecol Cancer Soc. 2003 Dec;13 Suppl 2:144–8.

34. Pepa CD, Tonini G, Pisano C, Di Napoli M, Cecere SC, Tambaro R, et al. Ovarian cancer standard of care: are there real alternatives? Chin J Cancer. 2015 Jan;34(1):17–27.

35. Swenerton K, Jeffrey J, Stuart G, Roy M, Krepart G, Carmichael J, et al. Cisplatin-cyclophosphamide versus carboplatin-cyclophosphamide in advanced ovarian cancer: a randomized phase III study of the National Cancer Institute of Canada Clinical Trials Group. J Clin Oncol Off J Am Soc Clin Oncol. 1992 May;10(5):718–26.

36. du Bois A, Lück H-J, Meier W, Adams H-P, Möbus V, Costa S, et al. A randomized clinical trial of cisplatin/paclitaxel versus carboplatin/paclitaxel as first-line treatment of ovarian cancer. J Natl Cancer Inst. 2003 Sep 3;95(17):1320–9.

37. Perren TJ, Swart AM, Pfisterer J, Ledermann JA, Pujade-Lauraine E, Kristensen G, et al. A phase 3 trial of bevacizumab in ovarian cancer. N Engl J Med. 2011 Dec 29;365(26):2484–96.

38. Burger RA, Fleming GF, Mannel RS, Greer BE, Liang SX. Incorporation of Bevacizumab in the Primary Treatment of Ovarian Cancer. N Engl J Med. 2011;11.

39. Schwartz LH, Litière S, de Vries E, Ford R, Gwyther S, Mandrekar S, et al. RECIST 1.1 – Update and Clarification: From the RECIST Committee. Eur J Cancer Oxf Engl 1990. 2016 Jul;62:132–7.

40. Moore K, Colombo N, Scambia G, Kim B-G, Oaknin A, Friedlander M, et al. Maintenance Olaparib in Patients with Newly Diagnosed Advanced Ovarian Cancer. N Engl J Med. 2018 Dec 27;379(26):2495–505.

41. Tapia G, Diaz-Padill I. Molecular Mechanisms of Platinum Resistance in Ovarian Cancer. In: Diaz-Padilla I, editor. Ovarian Cancer - A Clinical and Translational Update [Internet]. InTech; 2013 [cited 2020 May 28]. Available from: http://www.intechopen.com/books/ovarian-cancer-a-clinical-and-translational-update/molecular-mechanisms-of-platinum-resistance-in-ovarian-cancer

42. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). Eur J Cancer. 2009 Jan;45(2):228–47.

43. Pujade-Lauraine E, Wagner U, Aavall-Lundqvist E, Gebski V, Heywood M, Vasey PA, et al. Pegylated liposomal Doxorubicin and Carboplatin compared with Paclitaxel and Carboplatin for patients with platinum-sensitive ovarian cancer in late relapse. J Clin Oncol Off J Am Soc Clin Oncol. 2010 Jul 10;28(20):3323–9.

44. Coleman RL, Oza AM, Lorusso D, Aghajanian C, Oaknin A, Dean A, et al. Rucaparib maintenance treatment for recurrent ovarian carcinoma after response to platinum therapy (ARIEL3): a randomised, double-blind, placebo-controlled, phase 3 trial. The Lancet. 2017 Oct 28;390(10106):1949–61.

45. Aghajanian C, Blank SV, Goff BA, Judson PL, Teneriello MG, Husain A, et al. OCEANS: A Randomized, Double-Blind, Placebo-Controlled Phase III Trial of Chemotherapy With or Without Bevacizumab in Patients With Platinum-Sensitive Recurrent Epithelial Ovarian, Primary Peritoneal, or Fallopian Tube Cancer. J Clin Oncol. 2012 Jun 10;30(17):2039–45.

46. Lord CJ, Ashworth A. PARP inhibitors: Synthetic lethality in the clinic. Science. 2017 Mar 17;355(6330):1152–8.

47. Mirza MR, Monk BJ, Herrstedt J, Oza AM, Mahner S, Redondo A, et al. Niraparib Maintenance Therapy in Platinum-Sensitive, Recurrent Ovarian Cancer. N Engl J Med. 2016 Dec 1;375(22):2154–64.

48. Pujade-Lauraine E, Hilpert F, Weber B, Reuss A, Poveda A, Kristensen G, et al. Bevacizumab combined with chemotherapy for platinum-resistant recurrent ovarian cancer: The AURELIA open-label randomized phase III trial. J Clin Oncol Off J Am Soc Clin Oncol. 2014 May 1;32(13):1302–8.

49. Cooke SL, Brenton JD. Evolution of platinum resistance in high-grade serous ovarian cancer. Lancet Oncol. 2011 Nov 1;12(12):1169–74.

50. Kelland L. The resurgence of platinum-based cancer chemotherapy. Nat Rev Cancer. 2007 Aug;7(8):573–84.

51. Rocha C, Silva M, Quinet A, Cabral-Neto J, Menck C. DNA repair pathways and cisplatin resistance: an intimate relationship. Clinics [Internet]. 2018 Oct 9 [cited 2020 Mar 10];73(Suppl 1). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6113849/?report=classic

52. Mantri Y, Lippard SJ, Baik M-H. Bifunctional Binding of Cisplatin to DNA: Why Does Cisplatin Form 1,2-Intrastrand Cross-links with AG, But Not with GA? J Am Chem Soc. 2007 Apr 25;129(16):5023–30.

53. Suchánková T, Kubíček K, Kašpárková J, Brabec V, Kozelka J. Platinum–DNA interstrand crosslinks: Molecular determinants of bending and unwinding of the double helix. J Inorg Biochem. 2012 Mar;108:69–79.

54. Deans AJ, West SC. DNA interstrand crosslink repair and cancer. Nat Rev Cancer. 2011 Jul;11(7):467-.

55. Ghosh S. Cisplatin: The first metal based anticancer drug. Bioorganic Chem. 2019 Jul;88:102925.

56. Damia G, Broggini M. Platinum Resistance in Ovarian Cancer: Role of DNA Repair. Cancers. 2019 Jan 20;11(1).

57. Kilari D, Guancial E, Kim ES. Role of copper transporters in platinum resistance. World J Clin Oncol. 2016 Feb 10;7(1):106–13.

58. Xu X, Duan L, Zhou B, Ma R, Zhou H, Liu Z. Genetic polymorphism of copper transporter protein 1 is related to platinum resistance in Chinese non-small cell lung carcinoma patients. Clin Exp Pharmacol Physiol. 2012 Sep;39(9):786–92.

59. Ishida S, McCormick F, Smith-McCune K, Hanahan D. Enhancing tumor-specific uptake of the anticancer drug cisplatin with a copper chelator. Cancer Cell. 2010 Jun 15;17(6):574–83.

60. Sun S, Cai J, Yang Q, Zhao S, Wang Z. The association between copper transporters and the prognosis of cancer patients undergoing chemotherapy: a meta-analysis of literatures and datasets. Oncotarget. 2016 Dec 12;8(9):16036–51.

61. Holzer AK, Howell SB. The internalization and degradation of human copper transporter 1 following cisplatin exposure. Cancer Res. 2006 Nov 15;66(22):10944–52.

62. Taniguchi K, Wada M, Kohno K, Nakamura T, Kawabe T, Kawakami M, et al. A Human Canalicular Multispecific Organic Anion Transporter (cMOAT) Gene Is Overexpressed in Cisplatin-resistant Human Cancer Cell Lines with Decreased Drug Accumulation. Cancer Res. 1996 Sep 15;56(18):4124–9.

63. Hinoshita E, Uchiumi T, Taguchi K, Kinukawa N, Tsuneyoshi M, Maehara Y, et al. Increased expression of an ATP-binding cassette superfamily transporter, multidrug resistance protein 2, in human colorectal carcinomas. Clin Cancer Res Off J Am Assoc Cancer Res. 2000 Jun;6(6):2401–7.

64. Yamasaki M, Makino T, Masuzawa T, Kurokawa Y, Miyata H, Takiguchi S, et al. Role of multidrug resistance protein 2 (MRP2) in chemoresistance and clinical outcome in oesophageal squamous cell carcinoma. Br J Cancer. 2011 Feb;104(4):707–13.

65. Korita PV, Wakai T, Shirai Y, Matsuda Y, Sakata J, Takamura M, et al. Multidrug resistance-associated protein 2 determines the efficacy of cisplatin in patients with hepatocellular carcinoma. Oncol Rep. 2010 Apr;23(4):965–72.

66. Luo Q, Wu X, Zhang Y, Shu T, Ding F, Chen H, et al. ARID1A ablation leads to multiple drug resistance in ovarian cancer via transcriptional activation of MRP2. Cancer Lett. 2018 Jul 28;427:9–17.

67. K N, A K, K O, K M, N N, Y T. Copper-transporting P-type Adenosine Triphosphatase (ATP7B) as a Cisplatin Based Chemoresistance Marker in Ovarian Carcinoma: Comparative Analysis With Expression of MDR1, MRP1, MRP2, LRP and BCRP [Internet]. Vol. 101, International journal of cancer. Int J Cancer; 2002 [cited 2020 May 31]. Available from: https://pubmed.ncbi.nlm.nih.gov/12216079/

68. Li Y-Q, Yin J-Y, Liu Z-Q, Li X-P. Copper efflux transporters ATP7A and ATP7B: Novel biomarkers for platinum drug resistance and targets for therapy. IUBMB Life. 2018;70(3):183–91.

69. Amable L. Cisplatin resistance and opportunities for precision medicine. Pharmacol Res. 2016 Apr;106:27–36.

70. Surowiak P, Materna V, Kaplenko I, Spaczyński M, Dietel M, Lage H, et al. Augmented expression of metallothionein and glutathione S-transferase pi as unfavourable prognostic factors in cisplatin-treated ovarian cancer patients. Virchows Arch Int J Pathol. 2005 Sep;447(3):626–33.

71. Ceccaldi R, O'Connor KW, Mouw KW, Li AY, Matulonis UA, D'Andrea AD, et al. A unique subset of epithelial ovarian cancers with platinum sensitivity and PARP inhibitor resistance. Cancer Res. 2015 Feb 15;75(4):628–34.

72. Altaha R, Liang X, Yu JJ, Reed E. Excision repair cross complementing-group 1: gene expression and platinum resistance. Int J Mol Med. 2004 Dec;14(6):959–70.

73. Friboulet L, Postel-Vinay S, Sourisseau T, Adam J, Stoclin A, Ponsonnailles F, et al. ERCC1 function in nuclear excision and interstrand crosslink repair pathways is mediated exclusively by the ERCC1-202 isoform. Cell Cycle Georget Tex. 2013 Oct 15;12(20):3298–306.

74. Brown R, Hirst GL, Gallagher WM, McIlwrath AJ, Margison GP, van der Zee AG, et al. hMLH1 expression and cellular responses of ovarian tumour cells to treatment with cytotoxic anticancer agents. Oncogene. 1997 Jul 3;15(1):45–52.

75. Freimund AE, Beach JA, Christie EL, Bowtell DDL. Mechanisms of Drug Resistance in High-Grade Serous Ovarian Cancer. Hematol Clin. 2018 Dec 1;32(6):983–96.

76. Lord CJ, Ashworth A. BRCAness revisited. Nat Rev Cancer. 2016 Feb;16(2):110–20.

77. Sung P, Klein H. Mechanism of homologous recombination: mediators and helicases take on regulatory functions. Nat Rev Mol Cell Biol. 2006 Oct;7(10):739–50.

78. Konecny GE, Kristeleit RS. PARP inhibitors for BRCA1/2-mutated and sporadic ovarian cancer: current practice and future directions. Br J Cancer. 2016 Nov;115(10):1157–73.

79. Konstantinopoulos PA, Ceccaldi R, Shapiro GI, D'Andrea AD. Homologous Recombination Deficiency: Exploiting the Fundamental Vulnerability of Ovarian Cancer. Cancer Discov. 2015 Nov;5(11):1137–54.

80. Vencken PMLH, Kriege M, Hoogwerf D, Beugelink S, Burg MEL van der, Hooning MJ, et al. Chemosensitivity and outcome of BRCA1- and BRCA2-associated ovarian cancer patients after first-line chemotherapy compared with sporadic ovarian cancer patients. Ann Oncol. 2011 Jun 1;22(6):1346–52.

81. Peng G, Chun-Jen Lin C, Mo W, Dai H, Park Y-Y, Kim SM, et al. Genome-wide transcriptome profiling of homologous recombination DNA repair. Nat Commun. 2014 Feb 20;5(1):3361.

82. Abkevich V, Timms KM, Hennessy BT, Potter J, Carey MS, Meyer LA, et al. Patterns of genomic loss of heterozygosity predict homologous recombination repair defects in epithelial ovarian cancer. Br J Cancer. 2012 Nov 6;107(10):1776–82.

83. Weigelt B, Comino-Méndez I, de Bruijn I, Tian L, Meisel JL, García-Murillas I, et al. Diverse *BRCA1* and *BRCA2* Reversion Mutations in Circulating Cell-Free DNA of Therapy-Resistant Breast or Ovarian Cancer. Clin Cancer Res. 2017 Nov 1;23(21):6708–20.

84. Mayor P, Gay LM, Lele S, Elvin JA. BRCA1 reversion mutation acquired after treatment identified by liquid biopsy. Gynecol Oncol Rep. 2017 Jun 21;21:57–60.

85. Prieske K, Prieske S, Joosse SA, Trillsch F, Grimm D, Burandt E, et al. Loss of BRCA1 promotor hypermethylation in recurrent high-grade ovarian cancer. Oncotarget. 2017 Sep 15;8(47):83063–74.

86. Xu Y, Gao W, Zhang Y, Wu S, Liu Y, Deng X, et al. ABT737 reverses cisplatin resistance by targeting glucose metabolism of human ovarian cancer cells. Int J Oncol. 2018 Sep 1;53(3):1055–68.

87. Williams J, Lucas PC, Griffith KA, Choi M, Fogoros S, Hu YY, et al. Expression of Bcl-xL in ovarian carcinoma is associated with chemoresistance and recurrent disease. Gynecol Oncol. 2005 Feb 1;96(2):287–95.

88. Brotin E, Meryet-Figuière M, Simonin K, Duval RE, Villedieu M, Leroy-Dudal J, et al. Bcl-XL and MCL-1 constitute pertinent targets in ovarian carcinoma and their concomitant inhibition is sufficient to induce apoptosis. Int J Cancer. 2010;126(4):885–95.

89. Thibault B, Genre L, Le Naour A, Broca C, Mery E, Vuagniaux G, et al. DEBIO 1143, an IAP inhibitor, reverses carboplatin resistance in ovarian cancer cells and triggers apoptotic or necroptotic cell death. Sci Rep [Internet]. 2018 Dec 14 [cited 2020 Mar 26];8. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6294826/

90. Patch A-M, Christie EL, Etemadmoghadam D, Garsed DW, George J, Fereday S, et al. Whole-genome characterization of chemoresistant ovarian cancer. Nature. 2015 May 28;521(7553):489–94.

91. Yang D, Sun Y, Hu L, Zheng H, Ji P, Pecot CV, et al. Integrated analyses identify a master microRNA regulatory network for the mesenchymal subtype in serous ovarian cancer. Cancer Cell. 2013 Feb 11;23(2):186–99.

92. Marchini S, Fruscio R, Clivio L, Beltrame L, Porcu L, Nerini IF, et al. Resistance to platinum-based chemotherapy is associated with epithelial to mesenchymal transition in epithelial ovarian cancer. Eur J Cancer. 2013 Jan;49(2):520–30.

93. Parikh A, Lee C, Joseph P, Marchini S, Baccarini A, Kolev V, et al. microRNA-181a has a critical role in ovarian cancer progression through the regulation of the epithelial–mesenchymal transition. Nat Commun. 2014 May;5(1):2977.

94. Petrillo M, Zannoni GF, Beltrame L, Martinelli E, DiFeo A, Paracchini L, et al. Identification of high-grade serous ovarian cancer miRNA species associated with survival and drug response in patients receiving neoadjuvant chemotherapy: a retrospective longitudinal analysis using matched tumor biopsies. Ann Oncol Off J Eur Soc Med Oncol. 2016 Apr;27(4):625–34.

95. Bagnoli M, Canevari S, Califano D, Losito S, Maio MD, Raspagliesi F, et al. Development and validation of a microRNA-based signature (MiROvaR) to predict early relapse or progression of epithelial ovarian cancer: a cohort study. Lancet Oncol. 2016 Aug;17(8):1137–46.

96. Tomar T, Alkema NG, Schreuder L, Meersma GJ, de Meyer T, van Criekinge W, et al. Methylome analysis of extreme chemoresponsive patients identifies novel markers of platinum sensitivity in high-grade serous ovarian cancer. BMC Med. 2017 Jun 23;15(1):116.

97. Etemadmoghadam D, deFazio A, Beroukhim R, Mermel C, George J, Getz G, et al. Integrated Genome-Wide DNA Copy Number and Expression Analysis Identifies Distinct Mechanisms of Primary Chemoresistance in Ovarian Carcinomas. Clin Cancer Res. 2009 Feb 15;15(4):1417–27.

98. Pennington KP, Walsh T, Harrell MI, Lee MK, Pennil CC, Rendi MH, et al. Germline and Somatic Mutations in Homologous Recombination Genes Predict Platinum Response and Survival in Ovarian, Fallopian Tube, and Peritoneal Carcinomas. Clin Cancer Res. 2014 Feb 1;20(3):764–75.

99. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA. 2013 Nov 27;310(20):2191–4.

100. Ganzfried BF, Riester M, Haibe-Kains B, Risch T, Tyekucheva S, Jazic I, et al. curatedOvarianData: clinically annotated data for the ovarian cancer transcriptome. Database J Biol Databases Curation. 2013;2013:bat013.

101. Seela F, Budow S. Mismatch formation in solution and on DNA microarrays: how modified nucleosides can overcome shortcomings of imperfect hybridization caused by oligonucleotide composition and base pairing. Mol Biosyst. 2008 Mar;4(3):232–45.

102. D'Andrade PN, Fulmer-Smentek S. Agilent MicroRNA Microarray Profiling System. In: Fan J-B, editor. Next-Generation MicroRNA Expression Profiling Technology: Methods and Protocols

[Internet]. Totowa, NJ: Humana Press; 2012 [cited 2020 May 2]. p. 85–102. (Methods in Molecular Biology). Available from: https://doi.org/10.1007/978-1-61779-427-8_6

103. Gandolfo LC, Speed TP. RLE plots: Visualizing unwanted variation in high dimensional data. PLoS ONE [Internet]. 2018 Feb 5 [cited 2020 Apr 29];13(2). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5798764/

104. Risso D, Ngai J, Speed TP, Dudoit S. Normalization of RNA-seq data using factor analysis of control genes or samples. Nat Biotechnol. 2014 Sep;32(9):896–902.

105. Jacob L, Gagnon-Bartsch JA, Speed TP. Correcting gene expression data when neither the unwanted variation nor the factor of interest are observed. Biostat Oxf Engl. 2016 Jan;17(1):16–28.

106. Martini P, Sales G, Massa MS, Chiogna M, Romualdi C. Along signal paths: an empirical gene set approach exploiting pathway topology. Nucleic Acids Res. 2013 Jan;41(1):e19.

107. Calura E, Martini P, Sales G, Beltrame L, Chiorino G, D'Incalci M, et al. Wiring miRNAs to pathways: a topological approach to integrate miRNA and mRNA expression profiles. Nucleic Acids Res. 2014 Jun 17;42(11):e96–e96.

108. Sales G, Calura E, Cavalieri D, Romualdi C. graphite - a Bioconductor package to convert pathway topology to gene network. BMC Bioinformatics. 2012 Jan 31;13:20.

109. Lewis BP, Burge CB, Bartel DP. Conserved Seed Pairing, Often Flanked by Adenosines, Indicates that Thousands of Human Genes are MicroRNA Targets. Cell. 2005 Jan 14;120(1):15–20.

110. Calura E, Martini P, Sales G, Beltrame L, Chiorino G, D'Incalci M, et al. Wiring miRNAs to pathways: a topological approach to integrate miRNA and mRNA expression profiles. Nucleic Acids Res. 2014 Jun;42(11):e96.

111. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, et al. STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res. 2015 Jan;43(Database issue):D447-452.

112. Stark C, Breitkreutz B-J, Reguly T, Boucher L, Breitkreutz A, Tyers M. BioGRID: a general repository for interaction datasets. Nucleic Acids Res. 2006 Jan 1;34(Database issue):D535-539.

113. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003 Nov;13(11):2498–504.

114. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinforma Oxf Engl. 2010 Mar 1;26(5):589–95.

115. DePristo MA, Banks E, Poplin RE, Garimella KV, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet. 2011 May;43(5):491–8.

116. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GRS, Thormann A, et al. The Ensembl Variant Effect Predictor. Genome Biol. 2016 Jun 6;17(1):122.

117. Cibulskis K, Lawrence MS, Carter SL, Sivachenko A, Jaffe D, Sougnez C, et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. Nat Biotechnol. 2013 Mar;31(3):213–9.

118. Talevich E, Shain AH, Botton T, Bastian BC. CNVkit: Genome-Wide Copy Number Detection and Visualization from Targeted DNA Sequencing. PLoS Comput Biol. 2016 Apr;12(4):e1004873.

119. Lai Z, Markovets A, Ahdesmaki M, Chapman B, Hofmann O, McEwen R, et al. VarDict: a novel and versatile variant caller for next-generation sequencing in cancer research. Nucleic Acids Res. 2016 20;44(11):e108.

120. Paila U, Chapman BA, Kirchner R, Quinlan AR. GEMINI: integrative exploration of genetic variation and genome annotations. PLoS Comput Biol. 2013;9(7):e1003153.

121. Olshen AB, Venkatraman ES, Lucito R, Wigler M. Circular binary segmentation for the analysis of array-based DNA copy number data. Biostat Oxf Engl. 2004 Oct;5(4):557–72.

122. Riester M, Singh AP, Brannon AR, Yu K, Campbell CD, Chiang DY, et al. PureCN: copy number calling and SNV classification using targeted short read sequencing. Source Code Biol Med. 2016 Dec 15;11(1):13.

123. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative Genomics Viewer. Nat Biotechnol. 2011 Jan;29(1):24–6.

124. Telli ML, Timms KM, Reid J, Hennessy B, Mills GB, Jensen KC, et al. Homologous Recombination Deficiency (HRD) Score Predicts Response to Platinum-Containing Neoadjuvant Chemotherapy in Patients with Triple-Negative Breast Cancer. Clin Cancer Res Off J Am Assoc Cancer Res. 2016 01;22(15):3764–73.

125. Sztupinszki Z, Diossy M, Krzystanek M, Reiniger L, Csabai I, Favero F, et al. Migrating the SNP array-based homologous recombination deficiency measures to next generation sequencing data of breast cancer. Npj Breast Cancer. 2018 Jul 2;4(1):1–4.

126. Zhang Z, Theurkauf WE, Weng Z, Zamore PD. Strand-specific libraries for high throughput RNA sequencing (RNA-Seq) prepared without poly(A) selection. Silence. 2012 Dec 28;3:9.

127. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013 Jan;29(1):15–21.

128. Pertea et al. - 2015 - StringTie enables improved reconstruction of a tra.pdf [Internet]. [cited 2020 Apr 14]. Available from: https://www-nature-com.libezproxy.open.ac.uk/articles/nbt.3122.pdf

129. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics. 2011 Aug 4;12(1):323.

130. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010 Jan 1;26(1):139–40.

131. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods. 2017 Apr;14(4):417–9.

132. Weigelt et al. - 2017 - Diverse iBRCA1i and iBRCA2i Reversion Mu.pdf [Internet]. [cited 2020 Mar 25]. Available from: https://clincancerres.aacrjournals.org/content/clincanres/23/21/6708.full.pdf

133. Alkema NG, Wisman GBA, van der Zee AGJ, van Vugt MATM, de Jong S. Studying platinum sensitivity and resistance in high-grade serous ovarian cancer: Different models for different questions. Drug Resist Updat. 2016 Jan;24:55–69.

134. Leung EL, Fraser M, Fiscus RR, Tsang BK. Cisplatin alters nitric oxide synthase levels in human ovarian cancer cells: involvement in p53 regulation and cisplatin resistance. Br J Cancer. 2008 Jun;98(11):1803–9.

135. Kang H, Escudero-Esparza A, Douglas-Jones A, Mansel RE, Jiang WG. Transcript analyses of stromal cell derived factors (SDFs): SDF-2, SDF-4 and SDF-5 reveal a different pattern of expression and prognostic association in human breast cancer. Int J Oncol. 2009 Jul;35(1):205–11.

136. Vendrell E, Ribas M, Valls J, Solé X, Grau M, Moreno V, et al. Genomic and transcriptomic prognostic factors in R0 Dukes B and C colorectal cancer patients. Int J Oncol. 2007 May;30(5):1099–107.

137. Willis S, Villalobos VM, Gevaert O, Abramovitz M, Williams C, Sikic BI, et al. Single Gene Prognostic Biomarkers in Ovarian Cancer: A Meta-Analysis. PLOS ONE. 2016 Feb 17;11(2):e0149183.

138. Zhou L, Ercolano E, Ammoun S, Schmid MC, Barczyk MA, Hanemann CO. Merlin-Deficient Human Tumors Show Loss of Contact Inhibition and Activation of Wnt/β-Catenin Signaling Linked to the PDGFR/Src and Rac/PAK Pathways. Neoplasia N Y N. 2011 Dec;13(12):1101–12.

139. Rong R, Tang X, Gutmann DH, Ye K. Neurofibromatosis 2 (NF2) tumor suppressor merlin inhibits phosphatidylinositol 3-kinase through binding to PIKE-L. Proc Natl Acad Sci. 2004 Dec 28;101(52):18200–5.

140. Serrano I, McDonald PC, Lock F, Muller WJ, Dedhar S. Inactivation of the Hippo tumour suppressor pathway by integrin-linked kinase. Nat Commun. 2013 Dec;4(1):2976.

141. Zhang C, Li A, Li H, Peng K, Wei Q, Lin M, et al. PPP1R12A Copy Number Is Associated with Clinical Outcomes of Stage III CRC Receiving Oxaliplatin-Based Chemotherapy. Mediators Inflamm [Internet]. 2015 [cited 2020 Jul 3];2015. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4465766/

142. Dann RB, DeLoia JA, Timms KM, Zorn KK, Potter J, Flake DD, et al. BRCA1/2 mutations and expression: response to platinum chemotherapy in patients with advanced stage epithelial ovarian cancer. Gynecol Oncol. 2012 Jun;125(3):677–82.

143. Domchek SM, Aghajanian C, Shapira-Frommer R, Schmutzler RK, Audeh MW, Friedlander M, et al. Efficacy and safety of olaparib monotherapy in germline BRCA1/2 mutation carriers with advanced ovarian cancer and three or more lines of prior therapy. Gynecol Oncol. 2016 Feb;140(2):199–203.

144. The Cancer Genome Atlas Research Network - 2011 - Integrated genomic analyses of ovarian carcinoma.pdf [Internet]. [cited 2020 Mar 13]. Available from: https://www.nature.com/articles/nature10166.pdf

145. Australian Pancreatic Cancer Genome Initiative, ICGC Breast Cancer Consortium, ICGC MMML-Seq Consortium, ICGC PedBrain, Alexandrov LB, Nik-Zainal S, et al. Signatures of mutational processes in human cancer. Nature. 2013 Aug;500(7463):415–21.

146. Polak P, Kim J, Braunstein LZ, Karlic R, Haradhavala NJ, Tiao G, et al. A mutational signature reveals alterations underlying deficient homologous recombination repair in breast cancer. Nat Genet. 2017 Oct;49(10):1476–86.

147. Ray-Coquard I, Pautier P, Pignata S, Pérol D, González-Martín A, Berger R, et al. Olaparib plus Bevacizumab as First-Line Maintenance in Ovarian Cancer. N Engl J Med. 2019 Dec 19;381(25):2416–28.

148. González-Martín A, Pothuri B, Vergote I, DePont Christensen R, Graybill W, Mirza MR, et al. Niraparib in Patients with Newly Diagnosed Advanced Ovarian Cancer. N Engl J Med. 2019 Dec 19;381(25):2391–402.

149. Loibl S, Weber KE, Timms KM, Elkin EP, Hahnen E, Fasching PA, et al. Survival analysis of carboplatin added to an anthracycline/taxane-based neoadjuvant chemotherapy and HRD score as predictor of response—final results from GeparSixto. Ann Oncol. 2018 Dec 1;29(12):2341–7.

150. Meng Y, Chen C-W, Yung MMH, Sun W, Sun J, Li Z, et al. DUOXA1-mediated ROS production promotes cisplatin resistance by activating ATR-Chk1 pathway in ovarian cancer. Cancer Lett. 2018 Aug;428:104–16.

151. Samuel P, Pink RC, Caley DP, Currie JMS, Brooks SA, Carter DRF. Over-expression of miR-31 or loss of KCNMA1 leads to increased cisplatin resistance in ovarian cancer cells. Tumour Biol J Int Soc Oncodevelopmental Biol Med. 2016 Feb;37(2):2565–73.

152. Dobbin ZC, Katre AA, Steg AD, Erickson BK, Shah MM, Alvarez RD, et al. Using heterogeneity of the patient-derived xenograft model to identify the chemoresistant population in ovarian cancer. Oncotarget. 2014 Aug 20;5(18):8750–64.

153. Cohen S, Mosig R, Moshier E, Pereira E, Rahaman J, Prasad-Hayes M, et al. Interferon regulatory factor 1 is an independent predictor of platinum resistance and survival in high-grade serous ovarian carcinoma. Gynecol Oncol. 2014 Sep 1;134(3):591–8.

154. Song J, Zhang W, Wang S, Liu K, Song F, Ran L. A panel of 7 prognosis-related long non-coding RNAs to improve platinum-based chemoresistance prediction in ovarian cancer. Int J Oncol. 2018 Aug;53(2):866–76.

155. Deng L, Broaddus RR, McCampbell A, Shipley GL, Loose DS, Stancel GM, et al. Identification of a novel estrogen-regulated gene, EIG121, induced by hormone replacement therapy and differentially expressed in type I and type II endometrial cancer. Clin Cancer Res Off J Am Assoc Cancer Res. 2005 Dec 1;11(23):8258–64.

156. Estrella JS, Ma LT, Milton DR, Yao JC, Wang H, Rashid A, et al. Expression of estrogen-induced genes and estrogen receptor β in pancreatic neuroendocrine tumors: implications for targeted therapy. Pancreas. 2014 Oct;43(7):996–1002.

157. Expression of KIAA1324 in ovarian cancer - The Human Protein Atlas [Internet]. [cited 2020 Jul 10]. Available from: https://www.proteinatlas.org/ENSG00000116299-KIAA1324/pathology/ovarian+cancer

158. Liu R-Z, Garcia E, Glubrecht DD, Poon HY, Mackey JR, Godbout R. CRABP1 is associated with a poor prognosis in breast cancer: adding to the complexity of breast cancer cell response to retinoic acid. Mol Cancer. 2015 Jul 5;14:129.

159. Miyake T, Ueda Y, Matsuzaki S, Miyatake T, Yoshino K, Fujita M, et al. CRABP1-reduced expression is associated with poorer prognosis in serous and clear cell ovarian adenocarcinoma. J Cancer Res Clin Oncol. 2011 Apr;137(4):715–22.

160. Bhasin N, Maynard TM, Gallagher PA, LaMantia A-S. Mesenchymal/epithelial regulation of retinoic acid signaling in the olfactory placode. Dev Biol. 2003 Sep 1;261(1):82–98.

161. Parikh A, Lee C, Joseph P, Marchini S, Baccarini A, Kolev V, et al. microRNA-181a has a critical role in ovarian cancer progression through the regulation of the epithelial–mesenchymal transition. Nat Commun. 2014 May;5(1):2977.

162. Hirohashi Y, Torigoe T, Mariya T, Kochin V, Saito T, Sato N. HLA class I as a predictor of clinical prognosis and CTL infiltration as a predictor of chemosensitivity in ovarian cancer. Oncoimmunology [Internet]. 2015 Mar 23 [cited 2020 Jul 13];4(5). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4485746/

163. Shehata M, Mukherjee A, Deen S, Al-Attar A, Durrant LG, Chan S. Human leukocyte antigen class I expression is an independent prognostic factor in advanced ovarian cancer resistant to first-line platinum chemotherapy. Br J Cancer. 2009 Oct 20;101(8):1321–8.

164. Pimiento JM, Chen D-T, Centeno BA, Davis-Yadley AH, Husain K, Fulp WJ, et al. Annexin A8 Is a Prognostic Marker and Potential Therapeutic Target for Pancreatic Cancer. Pancreas. 2015 Jan;44(1):122–7.

165. Mancini M, Gariboldi MB, Taiana E, Bonzi MC, Craparotta I, Pagin M, et al. Co-targeting the IGF system and HIF-1 inhibits migration and invasion by (triple-negative) breast cancer cells. Br J Cancer. 2014 Jun;110(12):2865–73.

166. Li B, Tsao SW, Chan KW, Ludwig DL, Novosyadlyy R, Li YY, et al. Id1-induced IGF-II and its autocrine/endocrine promotion of esophageal cancer progression and chemoresistance--implications for IGF-II and IGF-IR-targeted therapy. Clin Cancer Res Off J Am Assoc Cancer Res. 2014 May 15;20(10):2651–62.

167. Shimizu T, Sugihara E, Yamaguchi-Iwai S, Tamaki S, Koyama Y, Kamel W, et al. IGF2 preserves osteosarcoma cell survival by creating an autophagic state of dormancy that protects cells against chemotherapeutic stress. Cancer Res. 2014 Nov 15;74(22):6531–41.

168. Brouwer-Visser J, Huang GS. IGF2 signaling and regulation in cancer. Cytokine Growth Factor Rev. 2015 Jun;26(3):371–7.

169. Brouwer-Visser J, Lee J, McCullagh K, Cossio MJ, Wang Y, Huang GS. Insulin-Like Growth Factor 2 Silencing Restores Taxol Sensitivity in Drug Resistant Ovarian Cancer. Hawkins SM, editor. PLoS ONE. 2014 Jun 16;9(6):e100165.

170. Kao S-H, Wu K-J, Lee W-H. Hypoxia, Epithelial-Mesenchymal Transition, and TET-Mediated Epigenetic Changes. J Clin Med. 2016 Feb 4;5(2):24.

171. Liu Z, Yao Y, Huang S, Li L, Jiang B, Guo H, et al. LINC00662 promotes gastric cancer cell growth by modulating the Hippo-YAP1 pathway. Biochem Biophys Res Commun. 2018 02;505(3):843–9.

172. Gong W, Su Y, Liu Y, Sun P, Wang X. Long non-coding RNA Linc00662 promotes cell invasion and contributes to cancer stem cell-like phenotypes in lung cancer cells. J Biochem (Tokyo). 2018 Dec 1;164(6):461–9.